

基于分组的两步 DEA 基准学习

姚佳, 李俊桥, 张青

(陆军工程大学, 江苏 南京 210007)

摘要:数据包络分析(data envelopment analysis, DEA)在为决策单元(decision making unit, DMU)评估效率水平的同时,可为其中的非有效单元提供消除低效的改进措施,即基准信息。但经典 DEA 模型为非有效单元提供的基准信息不易一步到位,缺乏对分组信息的充分利用。在依赖上下文的 DEA 框架内进行开发,提出了一种基于分组的两步 DEA 基准学习模型。模型使用加权 L1 范式衡量待评估单元与相应目标的接近程度。通过最小化实际点到 Pareto 有效边界的距离,为每一个决策单元在组内和全局的最佳实践前沿上分别设立单独基准,解决了在实践中目标点难以一步实现的问题,模型的结果可以视为针对最佳实践的长期改进策略。由于充分考虑了分组信息,模型能够反映给定基准过程中涉及的 DMU 周围环境,并增强了组内 DMU 在设立目标上的灵活性。该模型被用于评估西班牙公立大学的科研水平,通过对比实验验证了该模型的优势。

关键词:数据包络分析; 决策单元; 基准学习; 目标设定

中图分类号:C934 **DOI:**10.12018/j.issn.2097-0730.20220906001

Group-Based Two-Step Benchmark Study with DEA

YAO Jia, LI Junqiao, ZHANG Qing

(Army Engineering University of PLA, Nanjing 210007, China)

Abstract: When data envelopment analysis (DEA) is applied to assess the efficiency level of the decision making unit (DMU), it can provide improvement measures, i. e., benchmarking information, to eliminate inefficiencies for non-effective units in it. However, the benchmarking information provided by the classical DEA model for non-effective units is not easy to implement in one step, and lacks the full use of grouping information. This paper develops and proposes a group-based two-step benchmarking model under the context-dependent DEA framework. The model uses the weighted L1 paradigm to measure the proximity between the evaluated DMU and its corresponding target. By minimizing the distance from the actual point to the Pareto boundary, separate benchmarks are established for each DMU on the intra-group and global best practice frontier, which solves the problem that the target point is difficult to achieve in one step in practice. The results of the model can be viewed as a long-term improvement strategy aimed at best practice. In addition, due to the full consideration of grouping information, the model can reflect the surrounding environment of DMU involved in a given benchmarking process and provide flexibility for setting up targets in the group. Finally, the model was used to assess the level of research in Spanish public universities, demonstrating the advantages of this model through comparative experiments.

Key words: data envelopment analysis; decision making unit (DMU); benchmark study; target setting

收稿日期:2022-09-06

基金项目:军内科研项目(KYJBJQZL2203)。

第一作者:姚佳,讲师,主要研究网络软件开发工具。

通信作者:李俊桥,讲师,主要研究现代通信技术和网络设计开发。

数据包络分析(data envelopment analysis, DEA)是以线性规划为基础,对同类型的多输入、多输出决策单元(decision making unit, DMU)进行相对有效性评估的非参数系统分析方法^[1]。其中,同类型条件要求所有的DMU具有相同的输入和输出指标、相同的任务和目标,以及相同的外部环境^[2]。DEA根据各DMU的观测数据,判断各DMU的相对有效性,本质是判断DMU偏离生产前沿面的程度^[3]。因此,DEA除了为每个DMU提供效率值以外,还将每个低效DMU在生产前沿面上投影,提供可用于提升DMU效率的基准信息。这种基准信息使DEA具有超越其他评估方法的明显优势。当前,DEA已被证明是一种为生产过程中涉及的DMU提供基准学习的有用工具,并已在许多领域得到了较好的应用^[4-5]。

正如Cook等^[6]所指出的:在基准学习的情况下,DEA定义的高效DMU可能不一定形成生产前沿,而是导致最佳实践前沿。但部分学者^[7]认为:DEA模型用于基准学习时,通常设定不切实际的目标,因为按照DEA分析结果制定的改进计划与实际表现相差太远。为了解决这一问题,Aparicio等^[8]提出了从DMU到DEA有效边界距离最小化模型。该模型在经典的加法DEA模型基础上取得了显著进步,使DMU能够以更少努力达到技术有效,并广泛应用于之后的研究中^[9-11]。

在许多DEA应用中,被评估单元往往根据其所处的环境被划分为不同的组。这意味着同一组的DMU所经历的情况相似,而不同组之间DMU所经历的情况可能相差很大。为了解决这一特定场景下的评估问题,Cook等^[12-13]将基准学习通用框架应用于分组条件下的基准学习,假设每组DMU都有一个通用基准,并为每组建立了通用最佳实践。然而,对于一些效率低下的DMU,由于基准目标与实际生产的偏差过大,Cook等^[12-13]模型提供的最接近的基准目标在实践中仍然无法实现。此外,结合公共权重集(common set of weights, CSW)为组内DMU建立共同基准,一方面会导致某些观察到的输入/输出水平恶化^[14-15];另一方面还将标准DEA中使用的技术效率概念转化为整体效率,忽略了组内个体之间的差异。Ramón等^[16]将基准学习通用框架扩展为基于不同组最佳实践的“交叉基准学习”方法。目前关于这一特定场景下基准学习的研究仍然缺乏。

Kwon等^[17]指出:基准学习的最终目标是追求改进,这种改进倾向于以连续且逐步的方式发生。从管理角度看,设定渐进和可操作的目标可以激励组织成员作出实现绩效目标的承诺。因此,分步基

准学习已成为DEA研究的热点。现有的分步基准学习方法可以大致分为两类:基于Onion-DEA的基准学习和基于DMU聚类的基准学习。

基于Onion-DEA的基准学习通过设置一系列嵌套有效前沿层的方式提供序列基准。Seiford等^[18]将价值判断纳入进度度量扩展了基准学习模型,并用于评估32款打印机的吸引力。Bougnol等^[19]使用Onion-DEA完成了大学排名。Johnson等^[20]在上下文依赖的DEA方法中开发了一个DEA辅助招聘过程。Kwon等^[17]将DEA与反向传播神经网络结合,提出了一种创新的三阶段模型,以支持“更好的实践”基准学习,而不是传统的“最佳实践”基准学习。Ramón等^[7]提出了一种基于上下文依赖的两步DEA基准学习模型,使所有DMU到DEA有效边界的距离之和最小。但上述分步都是在不断消除有效DMU的原则下进行的,在应用场景中缺乏可解释性。基准学习产生的层次结构没有相应的实际意义。

基于DMU聚类的基准学习根据DMU的输入和输出对其进行分组,目的是为每个DMU找到最相似的单元或单元组,以便形成有效的联盟。Alves等^[21-22]和Faezy等^[23]提出了一种“事前”DMU聚类方法,先使用自组织映射(self-organizing map, SOM)将团队分成若干个同质簇,然后使用DEA对每个簇中的DMU进行评估。Lu等^[24]提出了一种“事后”DMU聚类方法,先使用改进的基于松弛变量度量法模型完成所有DMU的评估,然后根据评估分数和层次聚类结果逐步为每个DMU找到“学习基准”。这种基于DMU聚类的基准学习在一定程度上提供了分步的实际意义,但仍有不足之处。例如,模型为同一簇中不同DMU设置的目标相同的可能性更大,这就无法为每个DMU提供个性化建议。

综合上述分析,本文假设DMU根据其经验的相似度被划分为不同的组,旨在建立分组假设下的分步基准学习模型。

1 通用基准学习模型

假设有 n 个同质决策单元使用 m 项输入产生 s 项输出,用符号 $(X_j, Y_j), j=1, \dots, n$ 表示。其中, $X_j=(x_{1j}, x_{2j}, \dots, x_{mj}) \geq 0, X_j \neq 0; Y_j=(y_{1j}, y_{2j}, \dots, y_{sj}) \geq 0, Y_j \neq 0$ 。根据Banker^[25]提出的生产可能集的定义: $T = \{(X, Y) \mid X \text{ can produce } Y\}$,在满足凸性公理、无效性公理和规模收益可变假设条

件下, T 可以表示为 $T = \{(X, Y) \mid X \geq \sum_{j=1}^n \lambda_j X_j, Y \leq \sum_{j=1}^n \lambda_j Y_j, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0\}$ 。

通过使用加权 L1 范式最小化目标点到实际输入输出的距离, 模型(1)为给定决策单元 DMU_0 在 T 的 Pareto 有效边界上找到最接近的目标。

$$\text{Min } \|(X_0, Y_0) - (\hat{X}_0, \hat{Y}_0)\|_1^w$$

s. t.

$$\sum_{j \in E} \lambda_j^E x_{ij} = x_{i0} - s_{i0}^{E-} \quad i = 1, \dots, m \quad (1a)$$

$$\sum_{j \in E} \lambda_j^E y_{rj} = y_{r0} + s_{r0}^{E+} \quad r = 1, \dots, s \quad (1b)$$

$$\sum_{j \in E} \lambda_j^E = 1 \quad (1c)$$

$$-\sum_{i=1}^m v_i^E x_{ij} + \sum_{r=1}^s u_r^E y_{rj} + u_0^E + d_j^E = 0 \quad j \in E \quad (1d)$$

$$v_i^E x_{i0} \geq 1 \quad i = 1, \dots, m \quad (1e)$$

$$u_r^E y_{r0} \geq 1 \quad r = 1, \dots, s \quad (1f)$$

$$d_j^E \leq M b_j^E \quad j \in E \quad (1g)$$

$$\lambda_j^E \leq M(1 - b_j^E) \quad j \in E \quad (1h)$$

$$s_{i0}^{E-} \geq 0 \quad i = 1, \dots, m$$

$$s_{r0}^{E+} \geq 0 \quad r = 1, \dots, s$$

$$\lambda_j^E, d_j^E \geq 0 \quad b_j^E \in \{0, 1\} \quad j \in E$$

$$u_0^E \quad \text{free} \quad (1)$$

其中:

(1) 加权 L1 范式, 代表输入和输出松弛的加权和。

$$\|(X_0, Y_0) - (\hat{X}_0, \hat{Y}_0)\|_1^w =$$

$$\sum_{i=1}^m \frac{x_{i0} - \hat{x}_{i0}}{x_{i0}} + \sum_{r=1}^s \frac{\hat{y}_{r0} - y_{r0}}{y_{r0}} =$$

$$\sum_{i=1}^m \frac{s_{i0}^{E-}}{x_{i0}} + \sum_{r=1}^s \frac{s_{r0}^{E+}}{y_{r0}}$$

(2) M 是一个足够大的正实数。

(3) E 是 T 的最有效 DMU 集合^[26], 需要在求解模型(1)之前被确定。

模型(1)为每一个 DMU_j 分别设置独立基准 (\hat{X}_j, \hat{Y}_j) 。需要强调的是, 加权 L1 范式解决了经典 DEA 模型在投入产出之间可能存在的不可公度性问题^[8]。式(1a-1c)确保基准点 (\hat{X}_j, \hat{Y}_j) 位于生产可能集合 T 内; 式(1d)中, 非零权重 v_i^E, u_r^E 构成了 T 的支撑超平面系数; 式(1e-1f)是不变加法模型^[27]的对偶式, 确保模型(1)具有度量单位不变性; 式(1g-1h)是模型(1)中的关键限制, 通过连接前

两约束条件, 以确保基准点落在生产可能集 T 的 Pareto 有效边界上。若 $\lambda_j^E > 0$, 式(1h)表明 $b_j^E = 0$, 由式(1g)可以得到 $d_j^E = 0$ 。因此, 如果一个属于 E 的有效决策单元 DMU_j 是被评估单元 DMU_0 的参考点, 那么 DMU_j 对应的空间坐标必定在所有参考点组成的超平面 $-\sum_{i=1}^m v_i^E x_{ij} + \sum_{r=1}^s u_r^E y_{rj} + u_0^E = 0$ 上, 该超平面属于生产可能集 T 。

结合模型(1)的最优解, 被评估单元 DMU_0 的基准点由式(2)确定。

$$\hat{X}_0^* = X_0 - S_0^{E-*} \quad (= \sum_{j \in E} \lambda_j^{E*} X_j) \quad (2)$$

$$\hat{Y}_0^* = Y_0 + S_0^{E+*} \quad (= \sum_{j \in E} \lambda_j^{E*} Y_j)$$

式中: $S_0^{E-*} = (s_{10}^{E-*}, \dots, s_{m0}^{E-*})'$, $S_0^{E+*} = (s_{10}^{E+*}, \dots, s_{s0}^{E+*})'$ 。

2 基于分组的两步 DEA 基准学习

模型(1)给出了在不分组条件下的通用基准学习方案, 该方案同样适用于分组条件下的基准学习, 但在分组条件下使用该方案存在不能合理利用分组信息的弊端。针对分组的情况, Cook 等^[12]提出了一种“组内通用基准学习”方案, 该方案假定: 基准是由整套 DMU 确定的 Pareto 有效边界, 规划目标是最小化组内所有 DMU 到基准点的加权距离和。因此, 该方案可以有效利用分组的信息, 以组别为单位, 为组内 DMU 在 Pareto 有效边界上寻找共同基准, 最终同时投影组内的所有 DMU 到求得共同基准上。

然而, 该方案一方面不利于非有效单元改进, 因为在实际情况中, 非有效单元难以在一次评估周期内达到基准点的生产效果, 尤其当基准点投入和产出与实际生产能力之间的改动很大时, DMU 难以一步到位; 另一方面, 该方案实质是为同组的 DMU 求解一组 CSW, 并将同组单元投影到 Pareto 有效边界的由相同有效单元确定的单一超平面上, 存在部分观察指标一定程度的水平恶化问题^[14-15]。因此, 受嵌套有效前沿层思想的启发, 假定: 基准学习应该先组内学习再全局交流两步实现。

正如 Cook 等^[6]指出, 有效的 DMU 只能形成最佳实践前沿, 而不是生产前沿。即在 DEA 理论中, 生产前沿是未知且一定优于最佳实践前沿的, 但可以通过增加有效 DMU 的数量, 让最佳实践前沿不断逼近真实生产前沿。相反如果在使用 DEA 进行

评估并设置目标时,减少有效DMU的数量会导致评估的最佳实践前沿远离生产前沿,得到弱化的Pareto有效边界,规划模型求得最优解——输入和输出松弛加权和也会降低,最终为非有效DMU设置相对“宽松”的基准点,从而减小了非有效单元与基准点之间的步长。

结合以上讨论,基于分组的两步DEA基准学习将按照以下流程展开:

(1) 第一步基准学习仅在组内完成交流。假设 J 表示被评估单元 DMU_0 所在分组集合,首先需要确定分组 J 的组内有效单元的集合,由 J_E 表示。组内有效单元确定的组内最佳实践前沿显然弱于全局最佳实践前沿,因此得到的组内基准点相对“宽松”,对于非有效单元来说更具可操作性。

(2) 被评估单元 $DMU_0, 0 \in J$ 第一步基准学习结束后,改进的投入产出值 $(\hat{X}_0^J, \hat{Y}_0^J)$ 将用于第二步基准学习,这是为了确保基准学习的连续性。

(3) 第二步基准学习基于全局最佳实践前沿展开,因而第二步求得的基准点位于全局生产可能集 T 的Pareto有效边界上。需要强调的是,第一步基准学习后,改进的投入产出值 $(\hat{X}_0^J, \hat{Y}_0^J)$ 不能影响原有的全局最佳实践前沿,更不能改变最初全局有效单元的集合。即:在两步基准学习的过程中,整个评估环境的生产技术假定不变。

结合以上步骤,提供 DMU_0 所寻求目标序列的模型由式(3)给出。

$$\text{Min } \|(X_0, Y_0) - (\hat{X}_0^J, \hat{Y}_0^J)\|_1 \quad (3a)$$

$$\text{Min } \|(X_0^J, Y_0^J) - (\hat{X}_0, \hat{Y}_0)\|_1 \quad (3b)$$

s. t.

$$\begin{aligned} & \sum_{j \in J_E} \lambda_j^{J_E} x_{ij} = x_{i0} - s_{i0}^{J_E^-} \quad i = 1, \dots, m \\ & \sum_{j \in J_E} \lambda_j^{J_E} y_{rj} = y_{r0} + s_{r0}^{J_E^+} \quad r = 1, \dots, s \\ & \sum_{j \in J_E} \lambda_j^{J_E} = 1 \\ & - \sum_{i=1}^m v_i^{J_E} x_{ij} + \sum_{r=1}^s u_r^{J_E} y_{rj} + u_0^{J_E} + d_j^{J_E} = 0 \\ & \quad \quad \quad j \in J_E \\ & v_i^{J_E} x_{i0} \geq 1 \quad i = 1, \dots, m \\ & u_r^{J_E} y_{r0} \geq 1 \quad r = 1, \dots, s \\ & d_j^{J_E} \leq Mb_j^{J_E} \quad j \in J_E \\ & \lambda_j^{J_E} \leq M(1 - b_j^{J_E}) \quad j \in J_E \\ & \sum_{j \in E} \lambda_j^E x_{ij} = (x_{i0} - s_{i0}^{J_E^-}) - s_{i0}^{E^-} \quad i = 1, \dots, m \\ & \sum_{j \in E} \lambda_j^E y_{rj} = (y_{r0} + s_{r0}^{J_E^+}) + s_{r0}^{E^+} \quad r = 1, \dots, s \end{aligned}$$

$$\begin{aligned} & \sum_{j \in E} \lambda_j^E = 1 \\ & - \sum_{i=1}^m v_i^E x_{ij} + \sum_{r=1}^s u_r^E y_{rj} + u_0^E + d_j^E = 0 \quad j \in E \\ & v_i^E (x_{i0} - s_{i0}^{J_E^-}) \geq 1 \quad i = 1, \dots, m \\ & u_r^E (y_{r0} + s_{r0}^{J_E^+}) \geq 1 \quad r = 1, \dots, s \\ & d_j^E \leq Mb_j^E \quad j \in E \\ & \lambda_j^E \leq M(1 - b_j^E) \quad j \in E \\ & s_{i0}^{J_E^-} \geq 0 \quad s_{i0}^{E^-} \geq 0 \quad i = 1, \dots, m \\ & s_{r0}^{J_E^+} \geq 0 \quad s_{r0}^{E^+} \geq 0 \quad r = 1, \dots, s \\ & \lambda_j^{J_E}, d_j^{J_E} \geq 0 \quad b_j^{J_E} \in \{0, 1\} \quad j \in J_E \\ & u_0^{J_E} \quad u_0^E \quad \text{free} \\ & \lambda_j^E, d_j^E \geq 0 \quad b_j^E \in \{0, 1\} \quad j \in E \quad (3) \end{aligned}$$

式(3)是一个混合整数规划模型,同时也是一个多目标规划问题。为了简化模型计算,采用分层排序方法^[28],为式(3a)目标赋予更高的优先级。对于被评估单元 DMU_0 ,模型(3)的最优解提供了连续基准序列

$$\hat{X}_0^{J*} = X_0 - S_0^{J_E^-*} \quad (= \sum_{j \in J_E} \lambda_j^{J_E*} X_j) \quad (4)$$

$$\hat{Y}_0^{J*} = Y_0 + S_0^{J_E^+*} \quad (= \sum_{j \in J_E} \lambda_j^{J_E*} Y_j)$$

$$\hat{X}_0^* = (X_0 - S_0^{J_E^-*}) - S_0^{E^-*} \quad (= \sum_{j \in E} \lambda_j^{E*} X_j)$$

$$\hat{Y}_0^* = (Y_0 + S_0^{J_E^+*}) + S_0^{E^+*} \quad (= \sum_{j \in E} \lambda_j^{E*} Y_j) \quad (5)$$

式中: $S_0^{J_E^-*} = (s_{10}^{J_E^-*}, \dots, s_{m0}^{J_E^-*})'$, $S_0^{J_E^+*} = (s_{10}^{J_E^+*}, \dots, s_{s0}^{J_E^+*})'$ 。 $(\hat{X}_0^{J*}, \hat{Y}_0^{J*})$ 和 $(\hat{X}_0^*, \hat{Y}_0^*)$ 分别是 DMU_0 两次基准学习的投影点。

本文提出的模型充分考虑了分组的信息,为每一个DMU提供单独的基准点序列,该模型在目标设置时具有一定程度的灵活性。这意味着模型允许DMU被投影到Pareto边界的任何地方,每个单元可以设置自己的目标,只要该目标是有效的,最终由模型(3)找到这些目标中最接近的点。若一个有效 DMU_j 作为参考体,积极参与单元 DMU_0 的评估,那么该有效单元对应的空间坐标必定在最佳实践前沿的一个超平面上。由于模型(3)包含了 M 和二元变量,可以使用特殊有序集^[21](special ordered sets, SOS)重新定义约束来解决。该模型可用于但不限于两步基准学习的情况,当分组存在嵌套时,提出的方法可以直接扩展到考虑两个以上嵌套的DEA前沿,由此得到一个更长期的逐级学习的过程。

3 实例验证

为了验证提出的两步基准学习模型,将其应用于西班牙公立大学的研究活动评估。

西班牙公立大学可被视为一系列同质 DMU,它们开展类似活动并产生可比较的研究结果,因此可以为分析定义一组共同的输出。分析数据来自西班牙大学校长会议(conference of rectors of the Spanish universities, CRUE)报告。Ramón 等^[7]、Cook 等^[12]均在相关研究中使用过该公开数据集。同时,为了便于与 Cook 等^[12]方法对比,在指标选取上与 Cook 等^[12]保持一致。

由于大学研究的数据通常以索引指标的形式呈现,如发表的文章数或收入与学术人员之间的比率。Alves 等^[22]和 Cook 等^[6]为没有明确输入(without explicit input, WED)的 DEA 模型提供了理论框架。在这个框架基础上,将模型进行相应形式转换。

从侧重于基准和目标设定的角度对大学进行评估。样本由 42 个西班牙公立大学组成,按照地域分布分成 17 组。原始数据见文献^[12]。

3.1 模型结果

表 1 记录了西班牙公立大学研究活动的分组情况和本文模型的求解结果。最初的 VRS-DEA 分析显示,UAB、UAM、UCM、UCN、UPF 和 USC 6 所大学是技术有效的。对于 6 所技术有效的大学,在

本文模型的约束条件下,其参考集合由其本身构成,因此为它们设定的目标将与实际数据一致^[29]。注意到表 1 中有 10 个地区均只包含一所大学,对于这种组内只有一个 DMU 的情况,由于缺乏组内区分度,因此在第一步设置组内基准时,目标点与其本身实际数据一致。但这并不影响 10 所大学设立全局目标,如第 2 节所述,全局目标基于由整个单元集确定的 DEA 有效边界,在上述例子中,由本文模型求解的这 10 所大学的全局目标与通过使用单步基准(即模型(1))分析的结果相吻合,因此所提出的方法对于小尺寸的 DMU 组不存在敏感性问题。

观察表 1 的目标点可以发现,本文模型在设置组内基准时呈现明显的地域特性。CATALUÑA 在 Incomes 和 Articles 方面设定的目标明显高于其他地区,MADRID 在 Doctors 和 Thesis 方面表现格外突出,CASTILLA Y LEÓN 设置的 Incomes 目标最低。具体来说,按照模型建议的绩效模式作为改进方式,CATALUÑA 为 Incomes(平均每单位全日制学术人员 29 985 欧元)和 Articles(平均每位全日制学术人员发表超过 4.99 篇文章)设定了高目标,这些目标表明该组的 UDG 和 UDL 大学需要在第一步改进中针对这两个方面付出非常大的努力。同样位于 MADRID 的 UAH 和 UCAR 需要在第一步针对 Doctors 和 Thesis 付出巨大的努力。这再一次体现了两步基准学习在目标设定方面的优势。

表 1 目标设定

地区	大学	组内					全局				
		Doctors	Articles	Incomes	Thesis	Sexenios	Doctors	Articles	Incomes	Thesis	Sexenios
ANDALUCÍA	UAL	0.82	3.61	15 076.26	0.14	1.69	0.87	3.75	15 076.26	0.18	1.89
	UCA	0.83	3.55	14 503.80	0.14	1.69	0.88	3.73	14 503.80	0.18	1.89
	UCO	0.78	4.11	19 640.80	0.13	1.74	0.81	4.11	24 453.14	0.13	1.84
	UGR	0.83	3.50	14 034.57	0.14	1.68	0.88	3.72	14 034.57	0.18	1.89
	UHU	0.83	3.50	14 034.57	0.14	1.68	0.88	3.72	14 034.57	0.18	1.89
	UJA	0.83	3.50	14 034.57	0.14	1.68	0.88	3.72	14 034.57	0.18	1.89
	UMA	0.82	3.61	15 051.83	0.14	1.69	0.87	3.75	15 051.83	0.18	1.89
	UPO	0.82	3.58	14 792.43	0.14	1.69	0.88	3.74	14 792.43	0.18	1.89
CATALUÑA	USE	0.80	3.89	17 662.15	0.13	1.72	0.84	3.89	19 770.51	0.15	1.86
	UAB	0.75	6.15	25 944.33	0.23	2.14	0.75	6.15	25 944.33	0.23	2.14
	UBA	0.75	5.59	18 577.18	0.16	1.87	0.87	5.59	23 634.39	0.16	2.03
	UDG	0.75	5.59	18 577.18	0.16	1.87	0.87	5.59	23 634.39	0.16	2.03
	UDL	0.64	5.49	27 012.80	0.16	1.45	0.78	5.49	27 012.80	0.16	2.04
	UPF	0.63	4.13	29 159.89	0.20	2.08	0.63	4.13	29 159.89	0.20	2.08
CASTILLA Y LEÓN	URV	0.64	5.49	27 012.80	0.16	1.45	0.78	5.49	27 012.80	0.16	2.04
	UBU	0.65	2.16	8 965.08	0.09	1.27	0.89	3.67	12 526.56	0.19	1.90
	ULE	0.70	2.37	9 368.23	0.04	1.42	0.73	4.34	34 125.45	0.08	1.78
	USAL	0.67	2.55	6 174.08	0.14	1.64	0.89	3.67	12 526.56	0.19	1.90
COMUNIDAD VALENCIANA	UVA	0.65	2.16	8 965.08	0.09	1.27	0.89	3.67	12 526.56	0.19	1.90
	UA	0.67	4.67	14 297.09	0.12	1.67	0.88	4.67	15 516.97	0.19	2.03
	UJCS	0.65	2.98	17 355.79	0.07	1.30	0.73	4.34	34 125.45	0.08	1.78
VALENCIANA	UMH	0.68	5.01	13 685.35	0.13	1.75	0.88	5.01	16 526.56	0.19	2.07
	UVEG	0.68	5.35	13 073.61	0.14	1.82	0.87	5.35	17 536.14	0.19	2.12

续表

地区	大学	组内					全局				
		Doctors	Articles	Incomes	Thesis	Sexenios	Doctors	Articles	Incomes	Thesis	Sexenios
MADRID	UAH	0.89	3.79	12 879.32	0.19	1.92	0.89	3.79	12 879.32	0.19	1.92
	UAM	0.85	7.13	22 843.91	0.20	2.35	0.85	7.13	22 843.91	0.20	2.35
	UCAR	0.85	6.83	21 945.65	0.20	2.31	0.85	6.83	21 945.65	0.20	2.31
	UCM	0.89	3.67	12 526.56	0.19	1.90	0.89	3.67	12 526.56	0.19	1.90
GALICIA	ULC	0.64	4.19	18 789.25	0.14	1.17	0.83	4.19	22 906.69	0.14	1.85
	USC	0.86	5.47	30 794.68	0.11	1.85	0.86	5.47	30 794.68	0.11	1.85
	UVI	0.63	4.16	18 483.62	0.14	1.15	0.83	4.16	22 705.88	0.14	1.85
CANARIAS	ULL	0.80	2.74	12 982.92	0.09	1.24	0.89	3.68	12 982.92	0.19	1.90
	ULPGC	0.80	2.74	12 982.92	0.09	1.24	0.89	3.68	12 982.92	0.19	1.90
ARAGÓN	UZA	0.65	3.10	16 889.15	0.09	1.50	0.75	4.28	32 161.91	0.09	1.79
ASTURIAS	UOV	0.69	3.80	19 120.76	0.11	1.58	0.78	4.16	28 234.84	0.11	1.81
CANTABRIA	UCN	0.73	4.34	34 125.45	0.08	1.78	0.73	4.34	34 125.45	0.08	1.78
CASTILLA LA MANCHA	UCLM	0.62	2.25	17 629.98	0.10	1.05	0.86	3.83	17 629.98	0.16	1.87
EXTREMA-DURA	UEX	0.62	2.45	23 135.07	0.07	1.30	0.73	4.34	34 125.45	0.08	1.78
ISLAS BALEARES	UIB	0.61	4.04	14 388.48	0.07	1.51	0.73	4.34	34 125.45	0.08	1.78
MURCIA	UMU	0.60	3.26	11 065.30	0.10	1.65	0.89	3.67	12 526.56	0.19	1.90
NAVARRA	UPN	0.67	2.89	13 846.90	0.07	1.34	0.73	4.34	34 125.45	0.08	1.78
PAIS VASCO	UPV	0.62	2.74	18 327.81	0.10	1.23	0.76	4.22	30 198.38	0.10	1.80
LA RIOJA	URI	0.69	2.32	5 999.38	0.07	1.10	0.89	3.67	12 526.56	0.19	1.90

3.2 与 Cook's model 对比

为了进一步说明两步基准学习的优势,表2给出了不同模型求解后5个维度指标的平均值。

表2 不同模型求解后各指标的均值

模型	Doctors	Articles	Incomes	Thesis	Sexenios
Original data	0.67	3.28	15 929.83	0.11	1.45
Model(1)	0.83	4.21	20 852.60	0.16	1.90
Our model Step 1	0.73	3.92	17 041.82	0.13	1.61
Our model Step 2	0.83	4.41	21 426.41	0.16	1.92
Cook's model	0.86	4.24	18 890.74	0.17	1.91

经过对比可以得到,Our model Step 1 设立的目标小于 Model(1)和 Cook's model 运算的结果,大于 Original data 的值,这与上文关于不同模型基准讨论所预期的一样。以 UBU 为例,如果采用模型(1)制定改进计划,UBU 必须将全日制博士与全日制学术人员之间的比例从 0.481 提升到 0.893,还需同时改善其 Articles(平均每位全日制学术人员发表论文数从 1.57 篇提升到 3.67 篇)、Incomes(平均每单位全日制学术人员从 6 005.17 欧元提升到 12 526.56 欧元)、Thesis(比率从 0.09 提升到 0.19)、Sexenios(比值从 0.72 提升到 1.9)等指标,上升幅度要求均超过 100%。尽管模型(1)最大限度地缩短了与有效边界的距离,但如果大学想要在一个步骤中实现这些目标,就必须付出巨大的努力。对比本文模型提供的连续基准学习结果,这种方法设定了中间目标,允许该大学分配努力到两个阶段,

在更“宽松”的性能级别边界上产生目标,分配实现最佳实践性能所需的努力。因此,对于 UBU 而言,第一阶段仅需改善 Doctors 参数到 0.65、Articles 参数到 2.16 篇、Incomes 参数到 8 965.08 欧元、Sexenios 参数到 1.27,并无需调整 Thesis,大大缓解了同时大幅提升 5 项指标的压力。即:本文模型通过将目标分散,为非有效 DMU 提供的不再是一个单独的目标点,而是针对最佳实践的持续改进策略,更具可操作性。

3.3 与其他两步基准学习模型对比

为了更好地展示本文模型与其他两步基准学习模型的区别,又做了两个对比实验进一步说明本文模型的优势。如前所述,两步基准学习模型的实质是削弱最佳实践前沿,通过嵌套的有效前沿层实现分步目标。因此,在比较两步基准学习模型之间的差异时,只选择第一步目标进行比较分析,因为第一步目标是展示模型差异的关键。

首先,选取 Onion-DEA 模型作为第一个对比实验,该模型通过不断消除有效的 DMU 实现两步基准;其次,选取 Alves 等^[21-22]提出的 SOM-BCC 模型作为第二个对比实验,该模型采用 SOM 方法进行 DMU 聚类,并根据聚类结果完成簇内基准学习。采用 SOM 方法得到的 DMU 聚类结果如表 3 所示。

为了比较 3 个模型在目标设定上的差异,接下来以 CASTILLA Y LEÓN 为例进行结果分析,结

果如表4所示。

表3 SOM方法的聚类结果

No.	DMUs
1	UCA, UPO, USE, UDL, UPF, UCAR, UCN, UCLM, UEX, UPV
2	UHU, UMA, URV, UJCS, USC
3	UAL, UCO, UAB, UA, ULC, UVI, ULL, UZA, UOV, UPN
4	UGR, UBU, ULE, UVA, UAH, ULPGC, UIB
5	UJA, UBA, UDG, UAM, UCM, UMU
6	USAL, UMH, UVEG, URI

表4 不同模型为 CASTILLA Y LEÓN 设定的目标

大学	模型	Doctors	Articles	Incomes	Thesis	Sexenios
UBU	Our model	0.65	2.16	8 965.08	0.09	1.27
	Onion-DEA	0.48	2.17	21 794.37	0.11	1.87
	SOM-BCC	0.83	3.50	14 034.57	0.14	1.68
ULE	Our model	0.70	2.37	9 368.23	0.04	1.42
	Onion-DEA	0.70	2.99	16 988.99	0.13	1.75
	SOM-BCC	0.83	3.50	14 034.57	0.14	1.68
USAL	Our model	0.67	2.55	6 174.08	0.14	1.64
	Onion-DEA	0.83	3.50	14 034.57	0.14	1.68
	SOM-BCC	0.68	5.25	12 817.90	0.14	1.81
UVA	Our model	0.65	2.16	8 965.08	0.09	1.27
	Onion-DEA	0.65	2.83	17 945.62	0.13	1.77
	SOM-BCC	0.83	3.50	14 034.57	0.14	1.68

从表4可以看出,除了UBU的Doctors之外,本文模型为CASTILLA Y LEÓN设定的目标都是最低的,使各大学在整改的第一步更容易实现目标。图1展示了表4中所有大学各个指标的平均值,由于不同指标的数值差异较大,将原始数据调整为1,分别进行了数值缩放处理。图1的结果再次展示了本文模型在目标设定方面的优势。

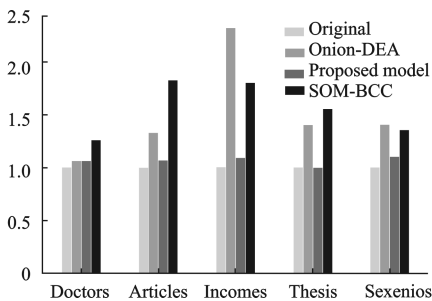


图1 不同模型为 CASTILLA Y LEÓN 设定的目标平均值

如前所述,本文模型在设定组内目标时具有明显的区域特征。然而,由于Onion-DEA和SOM-BCC没有考虑区域因素,这些隐藏的区域特征无法及时发现。一般认为地理上的分离增加了大学之间学习和交流的难度,同一地区的大学之间更易交流。同一地区大学之间的互助可以节约教育资源、共享成本,培养教育优势,形成区域教育特色。因此,本文模型在组内基准学习时只考虑同一地区大学之间的交流,以便更好地形成教育风格。此外,本

文模型也满足DEA同质性的要求。

除了将目标设置得太大和无法找到区域特征之外,使用SOM-BCC模型的另一个缺点是为同一簇内不同DMU设置的目标很可能是相同的。表5显示了第4个簇内的基准学习结果。可以看出,UGR,UBU,ULE,UVA,UAH,ULPGC在簇内共享同一个目标。产生这种现象的一个可能原因是SOM聚类时使用了索引信息,使每个簇内的DMU性能过于相似。同一簇内的不同DMU更易投射到同一最佳实践边界上的同一点。

表5 SOM-BCC模型第4个簇内的基准学习结果

大学	Doctors	Articles	Incomes	Thesis	Sexenios
UGR	0.83	3.50	14 034.57	0.14	1.68
UBU	0.83	3.50	14 034.57	0.14	1.68
ULE	0.83	3.50	14 034.57	0.14	1.68
UVA	0.83	3.50	14 034.57	0.14	1.68
UAH	0.83	3.50	14 034.57	0.14	1.68
ULPGC	0.83	3.50	14 034.57	0.14	1.68
UIB	0.61	4.04	14 388.48	0.07	1.51

4 结论

在Cook等^[12]的工作基础上,研究了分组条件下的基准学习模型。通过分析现有通用基准学习框架的弊端,采取为每个决策单元设立单独组内基准的方式;结合嵌套有效前沿层的思想,提出了基于分组的两步DEA基准学习模型,能够为低效单元提供连续且逐步的改进计划。该模型具有如下优势:(1)模型基于由整组DMU确定的DEA Pareto有效前沿设定全局目标,避免了小尺寸DMU组的敏感性问题;(2)由于充分考虑了分组信息,模型结果能更好地反映给定基准过程涉及的DMU周围环境,并允许对组内的DMU进行灵活处理;(3)模型在依赖于上下文的DEA框架内进行开发,将加权L1范式用作衡量评估单元的输入和输出与相应目标之间的接近程度,使用最小化到边界的距离作为优化目标,确保了每一步目标点的可接受性;(4)模型提供的基准点序列,可以视为针对最佳实践的长期改进策略。实验结果证明,这种方法对西班牙公立大学的研究绩效评估是有效的。

参考文献:

- [1] CHARNES A, COOPER W W, RHODES E. Measuring the efficiency of decision making units[J]. European Journal of Operational Research, 1978, 2(6): 429-444.
- [2] 杨国梁, 刘文斌, 郑海军. 数据包络分析方法(DEA)综述[J]. 系统工程学报, 2013, 28(6): 840-860. YANG Guoliang, LIU Wenbin, ZHENG Haijun. Re-

- view of data envelopment analysis[J]. *Journal of Systems Engineering*, 2013, 28(6): 840-860.
- [3] ZHU Q. Survey and prospect of data envelopment analysis (DEA) methods [J]. *Journal of Systems & Management*, 1994(4):1-9.
- [4] GUPTA P, MEHLAWAT M K, KUMAR A, et al. A credibilistic fuzzy DEA approach for portfolio efficiency evaluation and rebalancing toward benchmark portfolios using positive and negative returns[J]. *International Journal of Fuzzy Systems*, 2020, 22: 824-843.
- [5] JI X, WU J, ZHU Q Y, et al. Using a hybrid heterogeneous DEA method to benchmark China's sustainable urbanization: An empirical study[J]. *Annals of Operations Research*, 2019, 278(1): 281-335.
- [6] COOK W D, TONE K, ZHU J. Data envelopment analysis: Prior to choosing a model[J]. *Omega*, 2014, 44:1-4.
- [7] RAMÓN N, RUIZ J L, SIRVENT I. Two-step benchmarking: Setting more realistically achievable targets in DEA[J]. *Expert Systems With Applications*, 2018, 92: 124-131.
- [8] APARICIO J, RUIZ J L, SIRVENT I. Closest targets and minimum distance to the Pareto-efficient frontier in DEA[J]. *Journal of Productivity Analysis*, 2007, 28(3): 209-218.
- [9] RUIZ J L, SIRVENT I. Common benchmarking and ranking of units with DEA[J]. *Omega*, 2016, 65: 1-9.
- [10] LI F, ZHU Q Y, CHEN Z, et al. A balanced data envelopment analysis cross-efficiency evaluation approach [J]. *Expert Systems With Applications*, 2018, 106:154-168.
- [11] RUIZ J L, SIRVENT I. Performance evaluation through DEA benchmarking adjusted to goals[J]. *Omega*, 2019, 87: 150-157.
- [12] COOK W D, RUIZ J L, SIRVENT I, et al. Within-group common benchmarking using DEA[J]. *European Journal of Operational Research*, 2017, 256(3): 901-910.
- [13] COOK W D, ZHU J. Within-group common weights in DEA: An analysis of power plant efficiency[J]. *European Journal of Operational Research*, 2007, 178(1):207-216.
- [14] COOK W D, RAMÓN N, RUIZ J L, et al. DEA-based benchmarking for performance evaluation in pay-for-performance incentive plans [J]. *Omega*, 2019, 84: 45-54.
- [15] ROLL Y, COOK W D, GOLANY B. Controlling factor weights in data envelopment analysis [J]. *IIE Transactions*, 1991, 23(1): 2-9.
- [16] RAMÓN N, RUIZ J L, SIRVENT I. Cross-benchmarking for performance evaluation: Looking across best practices of different peer groups using DEA[J]. *Omega*, 2020, 92: 102169.
- [17] KWON H B, MARVEL J H, ROH J J. Three-stage performance modeling using DEA-BPNN for better practice benchmarking[J]. *Expert Systems with Applications*, 2017, 71: 429-441.
- [18] SEIFORD L M, ZHU J. Context-dependent data envelopment analysis-Measuring attractiveness and progress[J]. *Omega*, 2003, 31(5): 397-408.
- [19] BOUGNOL M L, DULÁJ H. Validating DEA as a ranking tool: An application of DEA to assess performance in higher education [J]. *Annals of Operations Research*, 2006, 145(1): 339-365.
- [20] JOHNSON S A, ZHU J. Identifying "best" applicants in recruiting using data envelopment analysis[J]. *Socio-Economic Planning Sciences*, 2003, 37(2):125-139.
- [21] ALVES L A, KRAMER H H, TSCHAFFON P B, et al. Assessing efficiency and setting benchmarks for NBA teams through DEA and DMU clustering[C]// *Proceedings of the 4th International Conference on Mathematics in Sport*. Leuven:[s. n.], 2013:1-7.
- [22] ALVES L A, MELLO J C C B S. Weights based clustering in data envelopment analysis using Kohonen neural network: An application in Brazilian electrical sector[J]. *IEEE Latin America Transactions*, 2015, 13(1): 188-194.
- [23] FAEZY R F, SHADLOO N. Portfolio selection using hybrid algorithm of data envelopment analysis based on Kohonen neural network and Cuckoo algorithm[J]. *Journal of Information and Optimization Sciences*, 2016, 37(4): 549-567.
- [24] LU W M, CHEN M H. A benchmark-learning roadmap for the military finance center [J]. *Mathematical and Computer Modelling*, 2011, 53(9/10): 1833-1843.
- [25] BANKER R D. Estimating most productive scale size using data envelopment analysis[J]. *European Journal of Operational Research*, 1984, 17(1): 35-44.
- [26] CHARNES A, COOPER W W, THRALL R M. A structure for classifying and characterizing efficiency and inefficiency in data envelopment analysis[J]. *Journal of Productivity Analysis*, 1991, 2(3): 197-237.
- [27] BAEK C, LEE J D. The relevance of DEA benchmarking information and the least-distance measure[J]. *Mathematical and Computer Modelling*, 2009, 49(1/2):265-275.
- [28] 陈亮, 陈雨果. 多目标发电调度策略与方法[J]. *电力系统自动化*, 2012, 36(9): 104-108.
CHEN Liang, CHEN Yuguo. Strategy and method for multi-objective generation dispatch schedules[J]. *Automation of Electric Power Systems*, 2012, 36(9): 104-108.
- [29] YANG Y S, GUO C Q, CHENG Y, et al. Research on production quality control and efficiency measurement based on DEA-benchmark model[J]. *Applied Mechanics and Materials*, 2012(190/191):151-155.