



# 计算机科学

COMPUTER SCIENCE

## 动态增量式逆向云变换算法研究

许昌林, 孔令卓

### 引用本文

许昌林, 孔令卓. 动态增量式逆向云变换算法研究[J]. 计算机科学, 2025, 52(5): 281-290.

XU Changlin, KONG Lingzhuo. [Research on Dynamic Incremental Backward Cloud Transformation Algorithm](#) [J]. Computer Science, 2025, 52(5): 281-290.

---

### 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

#### Similar articles recommended (Please use Firefox or IE to view the article)

#### [小样本飞机生产质量偏差数据分析与预测方法研究](#)

Study on Analysis and Prediction Method of Small Sample Aircraft Production Quality Deviation Data  
计算机科学, 2024, 51(11A): 240300123-8. <https://doi.org/10.11896/jsjcx.240300123>

#### [基于熵权-AHP与云模型的国产BIM建模软件多维度评价研究](#)

Multidimensional Evaluation Method for Domestic Building Information Modeling Software Based on Entropy-Weight-AHP and Cloud Model  
计算机科学, 2023, 50(6A): 220400216-9. <https://doi.org/10.11896/jsjcx.220400216>

#### [面向个性化需求的云制造服务可信评价模型](#)

Trust Evaluation Model of Cloud Manufacturing Services for Personalized Needs  
计算机科学, 2022, 49(3): 354-359. <https://doi.org/10.11896/jsjcx.210200116>

#### [基于改进樽海鞘算法的共享单车分布密度优化](#)

Optimization of Sharing Bicycle Density Distribution Based on Improved Salp Swarm Algorithm  
计算机科学, 2021, 48(11A): 106-110. <https://doi.org/10.11896/jsjcx.210700096>

#### [基于云模型和改进灰色关联分析模型的网络服务质量综合评估](#)

Comprehensive Evaluation of Network Service Quality Based on Cloud Model and Improved Grey Relational Analysis Model  
计算机科学, 2019, 46(5): 315-319. <https://doi.org/10.11896/j.issn.1002-137X.2019.05.049>

# 动态增量式逆向云变换算法研究

许昌林<sup>1,2</sup> 孔令卓<sup>1</sup>

1 北方民族大学数学与信息科学学院 银川 750021

2 北方民族大学宁夏智能信息与大数据处理重点实验室 银川 750021

**摘要** 云模型作为研究不确定性信息的工具,在不确定性人工智能和数据挖掘方面具有重要意义。其中逆向云变换算法为云模型的重要算法之一,可以实现定量数据到定性概念的转换。文中从动态增量的角度对逆向云变换算法进行研究。首先,针对现有的基于一阶绝对中心矩的经典逆向云变换算法中参数估计存在的不合理性进行了理论分析。其次,在理论分析的基础上,结合正向云变换算法生成云滴的特点,利用正态随机变量动态产生新的云滴作为新增样本,并将随机生成的样本和原有样本融合作为最终样本后再对参数进行估计,有效解决了已有算法存在的估算问题,从而提出了两种动态增量式的逆向云变换算法。然后,通过随机模拟实验,从有效性、稳定性、收敛性和参数的鲁棒性4个方面将所提出的逆向云变换算法与已有算法进行对比分析,实验结果表明所提出的动态增量式逆向云变换算法的估计误差较小、稳定性和收敛性较好,且对参数的变化具有较强的鲁棒性。最后,将提出的逆向云变换算法应用在对射击选手的射击水平模拟还原和评价中,实验结果进一步表明算法具有较好的实用性。

**关键词:** 云模型;逆向云变换;云滴;动态增量;超熵

**中图分类号** TP18

## Research on Dynamic Incremental Backward Cloud Transformation Algorithm

XU Changlin<sup>1,2</sup> and KONG Lingzhuo<sup>1</sup>

1 School of Mathematics and Information Science, North Minzu University, Yinchuan 750021, China

2 The Key Laboratory of Intelligent Information and Big Data Processing of NingXia Province, North Minzu University, Yinchuan 750021, China

**Abstract** As a tool for studying uncertain information, cloud model is of great significance in uncertain artificial intelligence and data mining. The backward cloud transformation algorithm is one of the important algorithms of cloud model, which can realize the transformation from quantitative data to qualitative concepts. This paper mainly studies the backward cloud transformation algorithm from the perspective of dynamic increment. Firstly, the irrationality of parameter estimation in the existing classical backward cloud transformation algorithm based on the first-order absolute central moment is analyzed theoretically. Secondly, on the basis of theoretical analysis, combined with the characteristics of cloud droplets generated by the forward cloud transformation algorithm, the normal random variable is used to dynamically generate new cloud droplets as new samples, then the randomly generated samples and the original samples are fused as the final samples to estimate the parameters, which effectively solves the estimation problems existing in the existing algorithms. Therefore, two dynamic incremental backward cloud transformation algorithms are proposed. Thirdly, through random simulation experiments, this paper compares the proposed backward cloud transform algorithm with existing algorithms from four aspects: effectiveness, stability, convergence and parameter robustness. The experimental results show that the dynamic incremental backward cloud transformation algorithm proposed in this paper has smaller estimation error, better stability and convergence, and has strong robustness to parameter changes. Finally, the proposed backward cloud transform algorithm is applied to the simulation and evaluation of Shooters' shooting level. The experimental results further show that the proposed algorithms have preferably practicability.

**Keywords** Cloud model, Backward cloud transformation, Cloud drop, Dynamic increment, Hyper entropy

到稿日期:2024-01-02 返修日期:2024-05-21

基金项目:宁夏自然科学基金优秀青年项目(2023AAC05046);国家自然科学基金(62066001);宁夏高等教育一流学科建设基金(NXYLXK2017B09)

This work was supported by the Excellent Youth Program of Ningxia Natural Science Foundation(2023AAC05046), National Natural Science Foundation of China(62066001) and Ningxia Higher Education First-class Discipline Construction Project(NXYLXK2017B09).

通信作者:许昌林(xuchlin@163.com)

## 1 引言

语言不确定性研究是近年来人工智能领域的重要工作之一<sup>[1-3]</sup>。随机性和模糊性是自然语言最基本的确定性<sup>[4-5]</sup>。概率理论从随机性角度将“随机性”用“概率”予以量化；模糊集和粗糙集理论从模糊性、不完备性角度对不确定性问题展开研究<sup>[6-8]</sup>；云模型理论从随机性、模糊性及两者关联性角度对不确定性展开研究<sup>[9-10]</sup>，实现了不确定性语言值的定量表示以及定性概念和定量数据间的不确定性转换，推动了不确定性问题的研究<sup>[11-12]</sup>。云模型在不同领域得到了较好应用<sup>[13-22]</sup>，例如 Jiang 等引入 G1-云模型对定性指标进行分析<sup>[13]</sup>，并将云模型应用于重庆长江上游航运中心建设成效评价中，取得了较好的评价效果；Ma 等将随机层次分析方法与云模型相结合，构建了新的科技投入项目绩效评估方法，并将其应用于科技投入项目绩效评估中<sup>[14]</sup>，从而提升了评估结果的科学性和合理性；Liu 等利用云模型理论刻画随机性和模糊性的特点<sup>[15]</sup>，构建了基于改进赋权法-云模型的地铁站突发事件应急管理评价模型，为地铁站应急管理能力的提升提供了理论基础；Xu 等基于正态云特征曲线并融合正态云分布特性，借鉴 Hellinger 距离刻画分布间相似性特点，提出了一种正态云相似性度量方法<sup>[16]</sup>，并将其应用于时间序列分类和协同过滤推荐系统中，取得了较好的效果；Liu 等针对专家给出的属性值为模糊语言且专家权重和属性权重均未知的多属性决策问题，构建了基于云模型的多属性决策方法<sup>[17-19]</sup>，较好地解决了决策信息的不确定性和决策过程的随机性问题，从而优化了复杂环境下的决策问题；Liu 等利用云模型研究不确定对象的聚类问题，提出了云聚类算法<sup>[20,22]</sup>，从而提高了平均聚类精度。

云模型利用期望(Expected value,  $Ex$ )、熵(Entropy,  $En$ )和超熵(Hyper entropy,  $He$ )等参数对不确定性概念内涵予以定量描述，通过正向云和逆向云变换算法实现概念内涵和概念外延间的互相转换<sup>[22-23]</sup>，其中逆向云变换算法是云模型应用于各个领域的重要环节。最早由 Liu 等提出的逆向云变换算法计算过程简便<sup>[24]</sup>，但该算法有时无法计算参数  $He$  的估计值；Yu 等提出了基于正态分布区间数的逆向云变换算法，避免了估计  $He$  过程中的重复计算问题<sup>[25]</sup>，但该算法只适用于定性概念为区间数的情况；Xu 等提出了两种多步逆向云变换算法<sup>[26-27]</sup>，解决了参数  $He$  无法估计的问题，且稳定性好、精度高，但其分组方式会影响参数估计值。Chen 等通过删除有效样本从而扩大样本方差的方式解决了参数  $He$  无法估计的问题<sup>[28]</sup>，但删除有效样本会造成信息损失，从而导致估计误差较大。基于此，本文针对逆向云变换算法进行了研究。首先，通过理论分析发现超熵  $He$  估计异常的原因在于熵  $En$  估计存在不足。其次，将动态新增样本与原有样本融合作为最终样本后再次估计熵  $En$  和超熵  $He$ ，直到计算出超熵的估计值  $\hat{He}$ ，从而提出了修正熵  $En$  的动态增量式逆向云变换算法和修正样本方差  $S^2$  的动态增量式逆向云变换算法。然后，通过随机模拟实验将提出的两种动态增量式逆向云变换算法与已有逆向云变换算法进行对比分析，结果表明所提出的方

法能有效解决超熵  $He$  无法估计的问题，且估计误差较小、稳定性和收敛性都较好，对参数变化具有较强的鲁棒性。最后，为了说明所提算法的有效性和实用性，将其应用于射击选手射击水平模拟还原和评价。

## 2 云模型相关理论

### 2.1 基本概念

**定义 1**<sup>[29-30]</sup> 设  $C$  是定量论域  $U$  上的定性概念，若定量值为定性概念  $C$  的一次随机实现，并且  $x$  对于  $C$  的确定度  $\mu(x) \in [0, 1]$  是有稳定倾向的随机数，则  $x$  在论域  $U$  上的分布称为云，每个  $x$  称为一个云滴。

由定义 1，不同概率分布构成不同云模型。由数字特征  $(Ex, En, He)$  表示的具有普适性的二阶正态云定义如下。

**定义 2**<sup>[16,29]</sup> 设  $C$  是定量论域  $U$  上的定性概念， $C$  包含  $(Ex, En, He)$  3 个数字特征，若定量值  $x \in U$  是定性概念  $C$  的一次随机实现，即  $x$  满足  $x = R_N(Ex, |y|)$ ，其中  $y = R_N(En, He)$ ，并且  $x$  对  $C$  的确定度  $\mu(x) \in [0, 1]$  是有稳定倾向的随机数。

$$\mu(x) = \exp\left\{-\frac{(x-Ex)^2}{2y^2}\right\} \quad (1)$$

则所有云滴在论域  $U$  上的分布称为二阶正态云，其中  $x = R_N(Ex, |y|)$  表示以  $Ex$  为期望、 $|y|$  为标准差的正态随机数。

### 2.2 已有正态云变换算法

正向云变换(Forward Cloud Transformation, FCT)和逆向云变换(Backward Cloud Transformation, BCT)可以实现概念外延与内涵间的转换<sup>[16]</sup>。正向云变换实现概念内涵到外延的转换，如算法 1 所示。

#### 算法 1 FCT<sup>[16]</sup>

输入：描述定性概念  $C$  的数字特征  $(Ex, En, He)$ ，产生云滴个数  $N$   
 输出： $N$  个云滴的定量值和其确定度  $\mu(x)$   
 Step1: 产生以  $En$  为期望值、 $He$  为均方差的一个正态随机数  $En' = R_N(En, He)$ 。  
 Step2: 产生以  $Ex$  为期望值、 $|En'|$  为均方差的一个正态随机数  $x_i = R_N(Ex, |En'|)$ 。  
 Step3: 求出  $x_i$  的确定度为  $\mu(x_i) = \exp\left\{-\frac{(x_i-Ex)^2}{2(En')^2}\right\}$ 。  
 Step4:  $x_i$  成为论域中的一个云滴，确定度为  $\mu(x_i)$ 。  
 Step5: 重复步骤 1-步骤 4，直至产生满足要求的  $N$  个云滴。

算法 SBCT-1stM 计算过程简便，但计算超熵估计值  $\hat{He} = \sqrt{S^2 - En^2}$  存在不足，有时无法得到  $\hat{He}$ 。例如，取参数  $(Ex, En, He)$  分别为  $(20, 3, 0.1)$ ，由算法 FCT 模拟产生  $N = 10000$  个样本，再由算法 SBCT-1stM 计算  $S^2 - \hat{En}^2$ 。这样反复执行  $T = 100$  次，且每次样本数据都重新生成，结果如图 1 所示。图 1 中近一半实验结果都出现了  $S^2 - \hat{En}^2 < 0$  的情况，意味着算法 SBCT-1stM 对超熵  $He$  的估计是失效的。

逆向云变换算法实现概念外延到内涵的转换，即得到参数  $(Ex, En, He)$  的估计值。目前已有多种逆向云变换算法，其中使用率较高的一阶绝对中心距逆向云变换算法(Backward Cloud Transformation Algorithm Based on the 1st-order Absolute Center Moment, SBCT-1stM)如算法 2 所示。

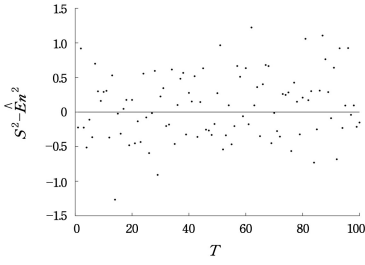


图1 算法 SBCT-1stM 计算 100 次时的  $S^2 - \hat{E}n^2$  值

Fig. 1 Algorithm SBCT-1stM calculates  $S^2 - \hat{E}n^2$  for 100 times

**算法 2** SBCT-1stM<sup>[23]</sup>

输入: N 个样本  $X_i (i=1, \dots, N)$

输出: 样本表征定性概念 C 的数字特征估计值  $(\hat{E}x, \hat{E}n, \hat{H}e)$

Step1: 根据给定的 N 个样本  $X_i$ , 计算样本均值  $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$ .

Step2:  $\hat{E}x = \bar{X}$ ,  $\hat{E}n = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_i - \hat{E}x|$ ,  $S^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \hat{E}x)^2$ .

Step3:  $\hat{H}e = \sqrt{S^2 - \hat{E}n^2}$ .

由云模型理论<sup>[31]</sup>, 二阶正态云 X 方差为  $DX = En^2 + He^2$ . 算法 SBCT-1stM 基于矩估计思想, 用样本方差  $S^2$  代替总体方差  $DX$ , 用  $\hat{E}n$  代替真实值  $En$ , 从而得到超熵  $\hat{H}e = \sqrt{S^2 - \hat{E}n^2}$ , 这个过程必然存在估计误差(理论分析见 3.1 节), 从而导致实际计算中出现  $S^2 - \hat{E}n^2 < 0$  的情况。具体表现为样本方差  $S^2$  偏小或熵的估计值  $\hat{E}n$  偏大。Chen 等通过逐步删除样本点从而增大样本方差  $S^2$  的方法来解决这个问题, 提出了一种逆向云变换算法(Stepwise Deleted Backward Cloud Transformation Algorithm, SDBCT)<sup>[28]</sup>。在算法 SD-BCT 中, 当云滴数  $N < 100$  时, 每次删除距离期望估计值  $\hat{E}x$  最近的 1 个样本点; 当给定的云滴数  $N \geq 100$ , 每次删除样本点总数的 1%。算法 SDBCT 解决了已有 SBCT-1stM 中  $\hat{H}e$  无法计算的问题, 但当样本量较大时, 需要删除的有效样本量也很大, 这必然增加估计值  $\hat{H}e$  的误差。

**3 动态增量式逆向云变换算法**

**3.1 算法提出的理论分析**

在云模型中超熵  $He$  是描述定性概念十分重要的数字特征, 影响着数据的分布。算法 SBCT-1stM 出现  $S^2 - \hat{E}n^2 < 0$  的根本原因在于估计式(2)存在不合理性。具体分析如下。

$$\hat{E}n = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_i - \hat{E}x| \quad (2)$$

由矩估计思想, 用样本一阶绝对中心矩估计总体一阶绝对中心矩, 即:

$$E|X - EX| = M|X - \bar{X}| \quad (3)$$

其中, 样本一阶绝对中心矩为:

$$M|X - \bar{X}| = \frac{1}{N} \sum_{i=1}^N |X_i - \bar{X}| \quad (4)$$

根据二阶正态云 X 的数学性质<sup>[31]</sup>, 有:

$$\begin{aligned} E|X - EX| &= \int_{-\infty}^{+\infty} |x - Ex| f_X(x) dx \\ &= \frac{1}{2\pi He} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |x - Ex| \frac{1}{|y|} e^{-\frac{(x-Ex)^2}{2y^2} - \frac{(y-En)^2}{2He^2}} dx dy \end{aligned} \quad (5)$$

其中,  $y = R_N(En, He)$  是正态随机数, 故对式(5)进行积分时需要分类讨论, 从而得到:

$$E|X - EX| = \begin{cases} \sqrt{\frac{\pi}{2}} En, & y > 0 \\ \sqrt{\frac{\pi}{2}} En + a, & y < 0 \end{cases} \quad (6)$$

其中,  $a = 2\sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{2\pi He}} \int_0^{+\infty} ye^{-\frac{(y+En)^2}{2He^2}} dy > 0$ 。

利用式(3)、式(4)和式(6), 可得到熵  $En$  的估计值  $\hat{E}n$  为:

$$\hat{E}n = \begin{cases} \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_i - \hat{E}x|, & y > 0 \\ \sqrt{\frac{\pi}{2}} \times \left( \frac{1}{N} \sum_{i=1}^N |X_i - \hat{E}x| - a \right), & y < 0 \end{cases} \quad (7)$$

不论  $y$  取何值, 算法 SBCT-1stM 对熵  $En$  的估计值都为:

$$\hat{E}n = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_i - \hat{E}x|$$

而由式(7)可知, 当  $y > 0$  时, 熵  $En$  的估计值为  $\hat{E}n = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_i - \hat{E}x|$ 。当  $y < 0$  时,

熵  $En$  的估计值为  $\hat{E}n = \sqrt{\frac{\pi}{2}} \times \left( \frac{1}{N} \sum_{i=1}^N |X_i - \hat{E}x| - a \right)$ , 其中

$a > 0$ 。因此, 当  $y < 0$  时, 算法 SBCT-1stM 对熵  $En$  的估计值比实际估计值大, 这是导致算法 SBCT-1stM 利用  $\hat{H}e =$

$\sqrt{S^2 - \hat{E}n^2}$  估计超熵  $He$  时出现估计值为虚数的根本原因。

综上所述, 由于  $a = 2\sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{2\pi He}} \int_0^{+\infty} ye^{-\frac{(y+En)^2}{2He^2}} dy$  无

解析解, 因此  $y < 0$  时, SBCT-1stM 得不到估计值  $\hat{E}n$  和  $\hat{H}e$ 。

逆向云变换是利用数据对参数  $(Ex, En, He)$  估计的过程, 本质上是参数估计过程<sup>[32]</sup>。理论上, 对样本信息利用越充分, 参数估计误差就越小。同样在逆向云变换中, 给定样本(云滴)越多, 参数  $(Ex, En, He)$  的估计值越精确。然而 SDBCT 在原有样本基础上逐步减少有效样本, 样本信息利用不充分导致估计误差增大。受此启发, 本文通过动态增加样本(云滴)修正熵  $En$  的估计和样本方差  $S^2$ , 从而提出两种动态增量式逆向云变换算法和修正方差的动态增量式逆向云变换算法。

修正熵的动态增量式逆向云变换算法

由 3.1 节分析可知, 当  $S^2 - \hat{E}n^2 < 0$  时, 需要对熵估计值  $\hat{E}n$  进行修正。本节修正策略是随机生成  $e \times N$  个以  $\hat{E}x$  为期望、以  $\lambda \times \hat{E}n (\lambda \leq 1)$  为方差的正态随机数作为新增样本( $e \in [0, 1]$ ), 将新增样本与原样本融合作为最终样本, 利用最终样

本修正熵  $En$  的估计和样本方差  $S^2$ , 从而提出两种动态增量式逆向云变换算法和修正方差的动态增量式逆向云变换算法。

修正熵的动态增量式逆向云变换算法

由 3.1 节分析可知, 当  $S^2 - \hat{E}n^2 < 0$  时, 需要对熵估计值  $\hat{E}n$  进行修正。本节修正策略是随机生成  $e \times N$  个以  $\hat{E}x$  为期望、以  $\lambda \times \hat{E}n (\lambda \leq 1)$  为方差的正态随机数作为新增样本( $e \in [0, 1]$ ), 将新增样本与原样本融合作为最终样本, 利用最终样

本修正熵  $En$  的估计和样本方差  $S^2$ , 从而提出两种动态增量式逆向云变换算法和修正方差的动态增量式逆向云变换算法。

修正熵的动态增量式逆向云变换算法

由 3.1 节分析可知, 当  $S^2 - \hat{E}n^2 < 0$  时, 需要对熵估计值  $\hat{E}n$  进行修正。本节修正策略是随机生成  $e \times N$  个以  $\hat{E}x$  为期望、以  $\lambda \times \hat{E}n (\lambda \leq 1)$  为方差的正态随机数作为新增样本( $e \in [0, 1]$ ), 将新增样本与原样本融合作为最终样本, 利用最终样

本再次估计熵  $En$  和超熵  $He$ , 直至  $S^2 - \hat{En}^2 > 0$ , 简记为 DIBCT( $\hat{En}$ ), 具体步骤如算法 3 所示, 流程图如图 2 所示。

**算法 3** DIBCT( $\hat{En}$ )

输入:  $N$  个样本  $X_i (i=1, \dots, N)$  和参数  $e$

输出: 样本表征定性概念的数字特征估计值 ( $\hat{Ex}, \hat{En}, \hat{He}$ )

Step1: 根据给定的  $N$  个样本  $X_i$ , 计算样本均值  $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$ , 并得到期望  $Ex$  的估计值为  $\hat{Ex} = \bar{X}$ 。

Step2: 计算样本方差  $S^2$  和熵  $En$  的估计值  $\hat{En}$ , 即

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \hat{Ex})^2, \hat{En} = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_i - \hat{Ex}|,$$

并令  $M = N$ 。

Step3: 令  $K = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_i - \hat{Ex}|$ 。

Step4: 如果  $S^2 - K^2 \geq 0$ , 则转向 Step6, 否则转向 Step5。

Step5: 生成  $e \cdot M$  个以  $\hat{Ex}$  为期望、 $\lambda_1 \cdot \hat{En} (\lambda_1 \leq 1)$  为方差的正态随机数作为新增样本, 并与原样本融合作为最终样本数据  $X_i$ , 此时样本量  $N$  修正  $N + e \cdot M$ , 再转向 Step3。

Step6:  $\hat{En} = K, \hat{He} = \sqrt{S^2 - \hat{En}^2}$ 。

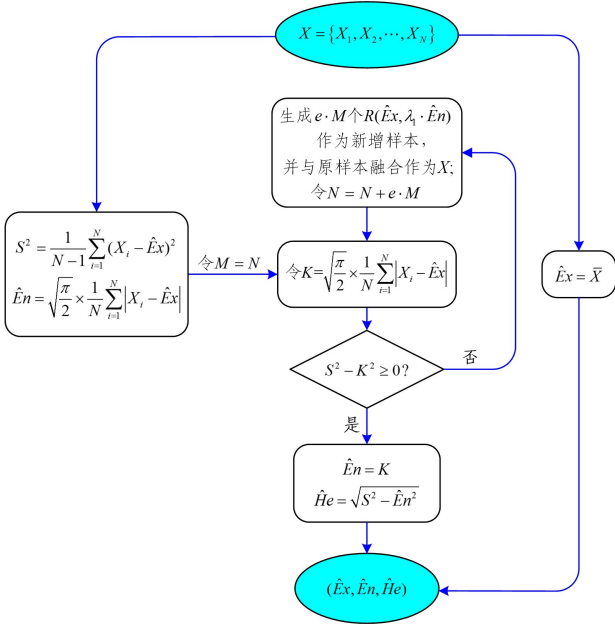


图 2 算法 DIBCT( $\hat{En}$ ) 的流程图

Fig. 2 Flowchart of algorithm DIBCT( $\hat{En}$ )

将算法 3 推广至高维情况, 得到修正熵的高维动态增量式逆向云变换算法 (High-dimensional Dynamic Incremental Backward Cloud Transformation Algorithm, HDIBCT( $\hat{En}$ )), 具体步骤如算法 4 所示。在 DIBCT( $\hat{En}$ ) 和 HDIBCT( $\hat{En}$ ) 中, 当条件  $S^2 - K^2 < 0$  和  $S_j^2 - K_j^2 < 0$  成立时, 意味着估计值  $\hat{En}$  偏大, 此时要产生正态随机数作为新增样本来修正  $\hat{En}$ , 因此  $\lambda_1 \times \hat{En}$  和  $\lambda_1 \cdot \hat{En}_j$  中参数  $\lambda_1 \leq 1$ , 可使产生的正态随机数融合到原有样本数据后能达到修正熵估计值  $\hat{En}$  的目的。

**算法 4** HDIBCT( $\hat{En}$ )

输入:  $N$  个  $m$  维样本  $X_i = \{x_{i1}, x_{i2}, \dots, x_{im}\}, (i=1, \dots, N)$  和参数  $e$   
 输出: 样本表征定性概念  $C$  的数字特征  $Ex = \{Ex_1, Ex_2, \dots, Ex_m\}$ ,  
 $En = \{En_1, En_2, \dots, En_m\}$  和  $He = \{He_1, He_2, \dots, He_m\}$  的估计值

Step1: 根据给定的  $N$  个样本  $X_i$ , 计算样本第  $j$  维样本均值  $\bar{X}_j = \frac{1}{N} \sum_{i=1}^N X_{ij}$ , 并得到  $\hat{Ex}_j = \bar{X}_j (j=1, 2, \dots, m)$ 。

Step2: 计算样本方差  $S_j^2$  和熵  $En_j$  的估计值  $\hat{En}_j$ , 分别为:

$$S_j^2 = \frac{1}{N-1} \sum_{i=1}^N (X_{ij} - \hat{Ex}_j)^2, \hat{En}_j = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_{ij} - \hat{Ex}_j|$$

并令  $M = N$ 。

Step3: 令  $K_j = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_{ij} - \hat{Ex}_j| (j=1, 2, \dots, m)$ 。

Step4: 如果  $S_j^2 - K_j^2 \geq 0$ , 则转向 Step6, 否则转向 Step5。

Step5: 生成  $e \cdot M$  个以  $\hat{Ex}_j$  为期望、 $\lambda_1 \cdot \hat{En}_j (\lambda_1 \leq 1)$  为方差的  $m$  维正态随机数作为新增样本, 并与原样本融合作为最终样本数据  $X_i$ , 此时样本量  $N$  修正为  $N + e \cdot M$ , 再转向 Step3。

Step6:  $\hat{En}_j = K_j, \hat{He}_j = \sqrt{S_j^2 - \hat{En}_j^2} (j=1, 2, \dots, m)$ 。

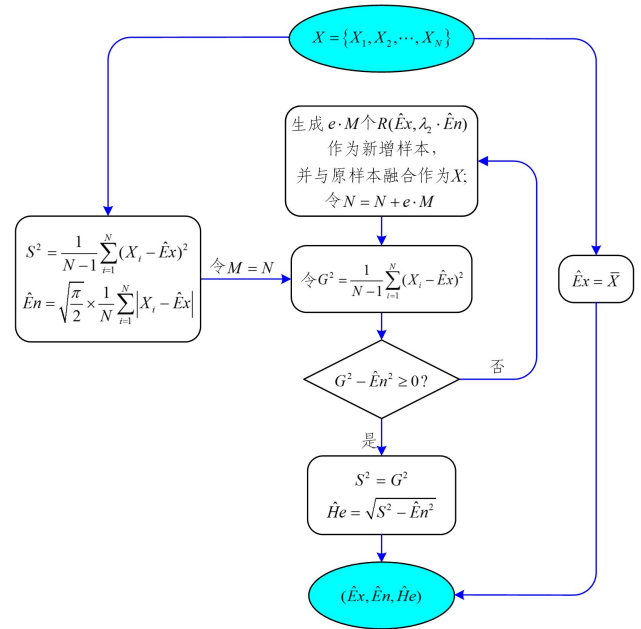


图 3 算法 DIBCT( $S^2$ ) 的流程图

Fig. 3 Flowchart of algorithm DIBCT( $S^2$ )

**3.3 修正方差的动态增量式逆向云变换算法**

通过实验发现, 用 3.2 节所述修正策略对样本方差  $S^2$  进行修正也能解决超熵估计值  $\hat{He}$  为虚数的问题, 因此本节给出另一种修正方差的动态增量式逆向云变换算法, 简记为 DIBCT( $S^2$ ), 具体步骤如算法 5 所示, 流程图如图 3 所示。进一步将算法 5 推广至高维情况, 得到修正方差的高维动态增量式逆向云变换算法, 记为 DIBCT( $S^2$ ), 如算法 6 所示。在 DIBCT( $S^2$ ) 和 HDIBCT( $S^2$ ) 中, 当条件  $G^2 - \hat{En}^2 < 0$  和  $G_j^2 - \hat{En}_j^2 < 0$  成立时, 意味着通过已有样本计算会出现  $S^2 < \hat{En}^2$ , 此时需产生正态随机数修正样本方差  $S^2$  使其变大, 因此  $\lambda_2 \times \hat{En}$  和  $\lambda_2 \times \hat{En}_j$  中的参数  $\lambda_2$  需大于等于 1, 从而使  $G^2 - \hat{En}^2 \geq 0$  和  $G_j^2 - \hat{En}_j^2 \geq 0$ 。

**算法 5** DIBCT( $S^2$ )

输入:  $N$  个样本  $X_i (i=1, \dots, N)$  和参数  $e$

输出: 样本表征定性概念  $C$  的数字特征估计值  $(\hat{E}_x, \hat{E}_n, \hat{H}_e)$

Step1: 根据给定  $N$  个样本  $X_i$ , 计算样本均值  $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$ , 并得到期望

$$E_x \text{ 估计值为 } \hat{E}_x = \bar{X}.$$

Step2: 计算样本方差  $S^2$  和熵  $E_n$  的估计值  $\hat{E}_n$ , 分别为:

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \hat{E}_x)^2, \hat{E}_n = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_i - \hat{E}_x|,$$

并令  $M=N$ .

Step3: 令  $G^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \hat{E}_x)^2$ .

Step4: 如果  $G^2 - \hat{E}_n^2 \geq 0$ , 则转向 Step6, 否则转向 Step5.

Step5: 生成  $e \cdot M$  个以  $\hat{E}_x$  为期望、 $\lambda_2 \cdot \hat{E}_n (\lambda_2 \geq 1)$  为方差的正态随机数作为新增样本, 并与原样本融合作为最终样本数据  $X_i$ , 此时样本量  $N$  修正为  $N+e \cdot M$ , 再转向 Step3.

Step6:  $S^2 = G^2, \hat{H}_e = \sqrt{S^2 - \hat{E}_n^2}$ .

**算法 6** HDIBCT( $S^2$ )

输入:  $N$  个  $m$  维样本  $X_i = \{x_{i1}, x_{i2}, \dots, x_{im}\} (i=1, \dots, N)$  和参数  $e$

输出: 样本表征定性概念  $C$  的数字特征  $E_x = \{E_{x1}, E_{x2}, \dots, E_{xm}\}$ , 因此

$E_n = \{E_{n1}, E_{n2}, \dots, E_{nm}\}$  和  $H_e = \{H_{e1}, H_{e2}, \dots, H_{em}\}$  的估计值

Step1: 根据给定的  $N$  个样本  $X_i$ , 计算样本第  $j$  维样本均值  $\bar{X}_j = \frac{1}{N} \sum_{i=1}^N X_{ij}$ , 并得到

$$\hat{E}_{xj} = \bar{X}_j (j=1, 2, \dots, m).$$

Step2: 计算样本方差  $S_j^2$  和熵  $E_{nj}$  的估计值  $\hat{E}_{nj}$ , 分别为:

$$S_j^2 = \frac{1}{N-1} \sum_{i=1}^N (X_{ij} - \hat{E}_{xj})^2, \hat{E}_{nj} = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |X_{ij} - \hat{E}_{xj}|$$

并令  $M=N$ .

Step3: 令  $G_j^2 = \frac{1}{N-1} \sum_{i=1}^N (X_{ij} - \hat{E}_{xj})^2 (j=1, 2, \dots, m)$ .

Step4: 如果  $G_j^2 - \hat{E}_{nj}^2 \geq 0$ , 则转向 Step6, 否则转向 Step5.

Step5: 生成  $e \cdot M$  个以  $\hat{E}_{xj}$  为期望、 $\lambda_2 \cdot \hat{E}_{nj} (\lambda_2 \geq 1)$  为方差的  $m$  维正态随机数作为新增样本, 并与原样本融合作为最终样本数据  $X_i$ , 此时样本量  $N$  修正为  $N+e \cdot M$ , 再转向 Step3.

Step6:  $S_j^2 = G_j^2, \hat{H}_{ej} = \sqrt{S_j^2 - \hat{E}_{nj}^2} (j=1, 2, \dots, m)$ .

在 DIBCT( $\hat{E}_n$ ), HDIBCT( $\hat{E}_n$ ), DIBCT( $S^2$ ) 和 HDIBCT( $S^2$ ) 中, 参数  $e$  的意义是在样本量  $N$  的基础上新增正态随机数的比例, 它是可调节的参数。这与算法 SDBCT 中的意义相同,  $e$  的改变会引起估计精度的改变(具体见模拟实验)。

此外, 与算法 SDBCT 相比, 本文给出的算法 DIBCT( $\hat{E}_n$ ), HDIBCT( $\hat{E}_n$ ), DIBCT( $S^2$ ) 和 HDIBCT( $S^2$ ) 由于充分利用了已有样本信息, 因此估计误差都较小(具体见模拟实验结果)。

**4 实验对比分析**

**4.1 实验说明**

本节利用随机模拟方式将提出算法 DIBCT( $\hat{E}_n$ ) 和 DIBCT( $S^2$ ) 与已有算法 SDBCT 进行对比分析, 即首先给定数字特征  $(E_x, E_n, H_e)$  真实值, 利用正向云变换算法 FCT 产生样本量为  $N$  的模拟数据, 然后分别利用算法 SDBCT,

DIBCT( $\hat{E}_n$ ) 和 DIBCT( $S^2$ ) 对数字特征  $(E_x, E_n, H_e)$  进行估计, 最后分别计算估计值的均值  $Mean(\hat{E}_n)$  与  $Mean(\hat{H}_e)$  和均方误差  $MSE(\hat{E}_n)$  与  $MSE(\hat{H}_e)$ , 即:

$$Mean(\hat{E}_n) = \frac{\sum_{k=1}^T \hat{E}_{nk}}{T}, Mean(\hat{H}_e) = \frac{\sum_{k=1}^T \hat{H}_{ek}}{T}$$

$$MSE(\hat{E}_n) = \frac{\sum_{k=1}^T (\hat{E}_{nk} - E_n)^2}{T}, MSE(\hat{H}_e) = \frac{\sum_{k=1}^T (\hat{H}_{ek} - H_e)^2}{T}$$

据此来分析 3 种逆向云变换算法的效果。

由于 SDBCT 和 DIBCT( $\hat{E}_n$ ) 以及 DIBCT( $S^2$ ) 对期望  $E_x$  的估计方法都相同, 因此没有对  $E_x$  的估计情况进行对比分析。同时, DIBCT( $\hat{E}_n$ ) 和 DIBCT( $S^2$ ) 估计熵  $E_n$  的方法相同, 因此也没有考虑 DIBCT( $S^2$ ) 对熵  $E_n$  的估计。此外, 从图 1 可以看出, 经典逆向云变换算法 SBCT-1stM 估计超熵  $H_e$  会出现虚数的异常情况, 因此后续实验未与算法 SBCT-1stM 进行对比分析。

算法 DIBCT( $\hat{E}_n$ ) 中  $\lambda_1 \cdot \hat{E}_n$  要求  $\lambda_1 \leq 1$ , 算法 DIBCT( $S^2$ )

中  $\lambda_2 \cdot \hat{E}_n$  要求  $\lambda_2 \geq 1$ , 因此实验统一取  $\lambda_1 = 0.98, \lambda_2 = 1.01$ 。结合文献[27, 31]对云概念的雾化分析和参数  $(E_x, E_n, H_e)$  的取值分析以及 2.2 节实验情况, 除特别说明外, 实验中统一取  $(E_x = 20, E_n = 3, H_e = 0.1)$ 。本节从参数  $e$  取定值时各种算法性能、稳定性和收敛性, 参数  $e$  动态增加时对各算法的影响, 以及超熵  $H_e$  和参数  $e$  同时变化时算法的性能等方面展开分析。

**4.2 参数  $e$  为定值时各算法的性能分析**

首先令  $(E_x = 20, E_n = 3, H_e = 0.1)$ , 由算法 FCT 产生容量为  $N = 10000$  的样本。其次, 利用 SDBCT, DIBCT( $\hat{E}_n$ ) 和 DIBCT( $S^2$ ) 分别估计  $E_n$  和  $H_e$ , 3 种算法中参数  $e = 0.01$ 。最后, 将上述实验过程重复  $T = 20$  次, 结果如图 4 和图 5 所示, 同时计算均值  $Mean(\hat{E}_n)$  和  $Mean(\hat{H}_e)$ 、均方误差  $MSE(\hat{E}_n)$  和  $MSE(\hat{H}_e)$ , 结果如表 1 和表 2 所列。

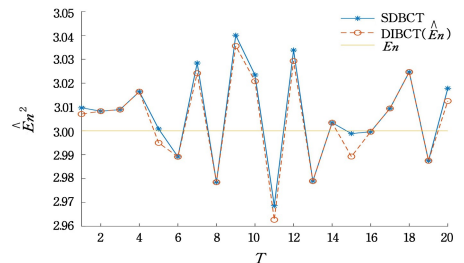


图 4 当  $T=20$  次时的估计值  $\hat{E}_n$

Fig. 4 Estimated value  $\hat{E}_n$  during  $T=20$  times

从图 4 可以看出, 在  $T = 20$  次的估计结果中, 算法 SDBCT 和 DIBCT( $\hat{E}_n$ ) 对熵的估计  $\hat{E}_n$  都在真值  $E_n = 3$  附近波动, 都能得到较好的估计值。同时, 从表 1 可以看出, 与算法 SDBCT 相比, 算法 DIBCT( $\hat{E}_n$ ) 的熵估计值的均值  $Mean(\hat{E}_n)$  更接近于真值, 均方误差  $MSE(\hat{E}_n)$  更小。这说明修正

熵  $E_n$  的逆向云变换算法  $DIBCT(\hat{E}_n)$  是有效的。

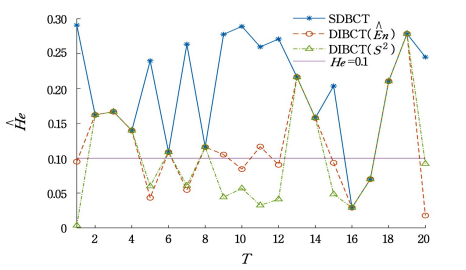


图 5 当  $T=20$  次时的估计值  $\hat{H}_e$

Fig. 5 Estimated value  $\hat{H}_e$  during  $T=20$  times

表 1 当  $T=20$  次时估计值  $\hat{E}_n$  的均值和均方误差

Table 1  $Mean(\hat{E}_n)$  and  $MSE(\hat{E}_n)$  when  $T=20$  times

均值与均方误差	SDBCT	$DIBCT(\hat{E}_n)$
$Mean(\hat{E}_n)$	3.0063	<b>3.0041</b>
$MSE(\hat{E}_n)$	$3.9029 \times 10^{-4}$	<b><math>3.5898 \times 10^{-4}</math></b>

表 2 当  $T=20$  时估计值  $\hat{H}_e$  的均值和均方误差

Table 2  $Mean(\hat{H}_e)$  and  $MSE(\hat{H}_e)$  when  $T=20$  times

均值与均方误差	SDBCT	$DIBCT(\hat{E}_n)$	$DIBCT(S^2)$
$Mean(\hat{H}_e)$	0.1998	<b>0.1178</b>	<b>0.1047</b>
$MSE(\hat{H}_e)$	0.0157	<b>0.0045</b>	<b>0.0053</b>

同样地,从图 5 可以看出,在  $T=20$  次超熵估计中,算法 SDBCT 由于会删除部分有效样本,因此估计值  $\hat{H}_e$  大多都大于真实值 0.1,而算法  $DIBCT(\hat{E}_n)$  和  $DIBCT(S^2)$  对超熵的估计值  $\hat{H}_e$  始终在真值 0.1 附近波动。进一步,由表 2 可知,  $DIBCT(\hat{E}_n)$  和  $DIBCT(S^2)$  对超熵的估计均值  $Mean(\hat{H}_e)$  都更接近真值 0.1,  $MSE(\hat{H}_e)$  都更小,这进一步说明  $DIBCT(\hat{E}_n)$  和  $DIBCT(S^2)$  对超熵  $H_e$  的估计是有效的且性能更好,同时也说明 SDBCT 将部分样本删除后再对  $H_e$  进行估计的做法会对  $H_e$  的估计结果产生较大的影响。

### 4.3 算法稳定性和收敛性分析

首先,令  $(Ex=20, En=3, He=0.1)$ , 由算法 FCT 分别产生容量为  $N=1000, 2000, \dots, 10000$  的随机样本。其次,对于每个  $N$ , 分别用 SDBCT,  $DIBCT(\hat{E}_n)$  和  $DIBCT(S^2)$  估计  $E_n$  和  $H_e$  (其中  $e=0.01$ ), 将实验重复  $T=100$  次, 计算  $Mean(\hat{E}_n)$  和  $Mean(\hat{H}_e)$ 、 $MSE(\hat{E}_n)$  和  $MSE(\hat{H}_e)$ 。不同样本量下熵  $E_n$  的估计值的  $Mean(\hat{E}_n)$  和  $MSE(\hat{E}_n)$  变化趋势图分别如图 6 和图 7 所示,超熵  $H_e$  估计值的  $Mean(\hat{H}_e)$  和  $MSE(\hat{H}_e)$  变化趋势分别如图 8 和图 9 所示。

从图 6 和图 7 可以看出,随着样本量增加,算法 SDBCT 和  $DIBCT(\hat{E}_n)$  对熵估计值的均值  $Mean(\hat{E}_n)$  都比较接近真值 3, 均方误差  $MSE(\hat{E}_n)$  整体呈现下降收敛趋势。与 SDBCT 相比,  $DIBCT(\hat{E}_n)$  对熵估计值的均方误差  $MSE(\hat{E}_n)$  收敛速度更快,这说明本文算法  $DIBCT(\hat{E}_n)$  对熵  $E_n$  的估计是稳定

且收敛的,而且效果更好。从图 8 可以看出,随着样本量的增加,与 SDBCT 相比,本文提出的两种算法  $DIBCT(\hat{E}_n)$  和  $DIBCT(S^2)$  对超熵估计值的均值  $Mean(\hat{H}_e)$  更接近真值 0.1。从图 9 可以看出,  $DIBCT(\hat{E}_n)$  和  $DIBCT(S^2)$  对超熵  $H_e$  估计的收敛速度要明显快于 SDBCT 的收敛速度,这说明  $DIBCT(\hat{E}_n)$  和  $DIBCT(S^2)$  对超熵  $H_e$  的估计是稳定且收敛的,而且效果比算法 SDBCT 更好。

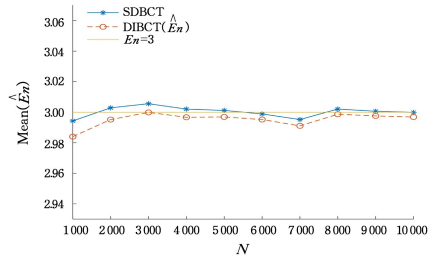


图 6 样本量增加时估计值  $\hat{E}_n$  的均值变化

Fig. 6  $Mean(\hat{E}_n)$  changes with increasing sample size

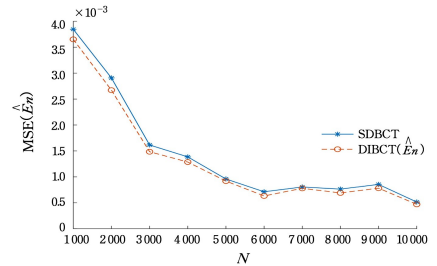


图 7 样本量增加时估计值  $\hat{E}_n$  的均方误差变化

Fig. 7  $MSE(\hat{E}_n)$  changes with increasing sample size

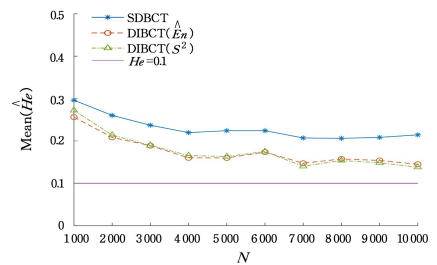


图 8 样本量增加时估计值  $\hat{H}_e$  的均值变化

Fig. 8  $Mean(\hat{H}_e)$  changes with increasing sample size

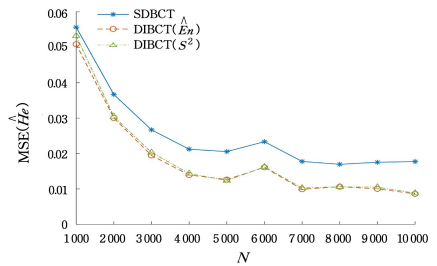


图 9 样本量增加时估计值  $\hat{H}_e$  的均方误差变化

Fig. 9  $MSE(\hat{H}_e)$  changes with increasing sample size

### 4.4 参数 $e$ 变化时对算法的影响分析

本节分析当参数  $e$  动态变化时,  $DIBCT(\hat{E}_n)$  和  $DIBCT$

( $S^2$ )与SDBCT对参数估计的适应情况。首先,给定( $Ex=20, En=3, He$ ),其中  $He=0.1$  和  $He=0.3$ ,利用算法 FCT 产生容量为  $N=10000$  的样本。其次,参数  $e$  以 0.01 为步长从 0.01 动态增加至 0.1,分别由 SDBCT,  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  估计  $En$  和  $He$ 。实验过程重复  $T=100$  次,分别计算  $T=100$  次时估计值的均值  $Mean(\hat{En})$  和  $Mean(\hat{He})$ 、均方差  $MSE(\hat{En})$  和  $MSE(\hat{He})$ 。当  $He=0.1$  时,  $Mean(\hat{En})$ ,  $Mean(\hat{He})$  和  $MSE(\hat{En})$ ,  $MSE(\hat{He})$  的变化情况分别如图 10—图 13 所示。当  $He=0.3$  时,均值  $Mean(\hat{En})$  和  $Mean(\hat{He})$ 、均方差  $MSE(\hat{En})$  和  $MSE(\hat{He})$  的变化情况分别如图 14—图 17 所示。

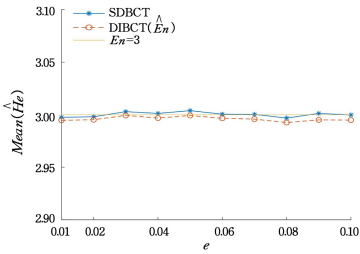


图 10  $e$  动态增加时估计值  $\hat{En}$  的均值 ( $He=0.1$ )

Fig. 10  $Mean(\hat{En})$  for  $e$  dynamically increasing ( $He=0.1$ )

不论  $He$  取 0.1 或 0.3,从图 10 和图 14 都可以看出,当  $e$  动态增加时,算法 SDBCT 和  $DIBCT(\hat{En})$  的均值  $Mean(\hat{En})$  都在真值 3 附近波动,且波动幅度小。进一步,从图 11 和图 15 可以看出,参数  $e$  动态增加时,算法 SDBCT 和  $DIBCT(\hat{En})$  的熵估计值的均方差  $MSE(\hat{En})$  都非常小,且不随  $e$  增大而大幅度变化。这表明在估计熵  $En$  时,算法 SDBCT 和  $DIBCT(\hat{En})$  对参数  $e$  的变化都不太敏感,都有较好的适应性。

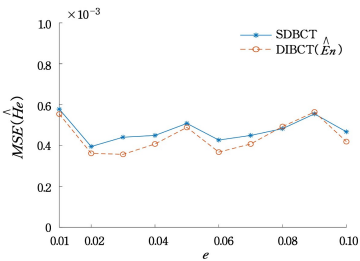


图 11  $e$  动态增加时估计值  $\hat{En}$  的均方差 ( $He=0.1$ )

Fig. 11  $MSE(\hat{En})$  for  $e$  dynamically increasing ( $He=0.1$ )

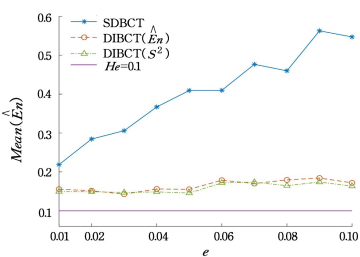


图 12  $e$  动态增加时估计值  $\hat{He}$  的均值 ( $He=0.1$ )

Fig. 12  $Mean(\hat{He})$  for  $e$  dynamically increasing ( $He=0.1$ )

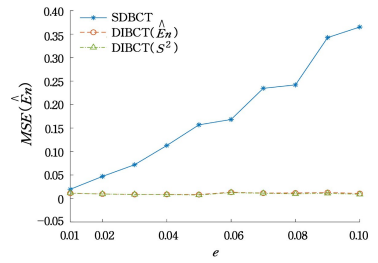


图 13  $e$  动态增加时估计值  $\hat{He}$  的均方差 ( $He=0.1$ )

Fig. 13  $MSE(\hat{He})$  for dynamically increasing ( $He=0.1$ )

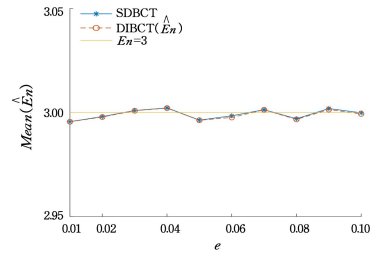


图 14  $e$  动态增加时估计值  $\hat{En}$  的均值 ( $He=0.3$ )

Fig. 14  $Mean(\hat{En})$  for  $e$  dynamically increasing ( $He=0.3$ )

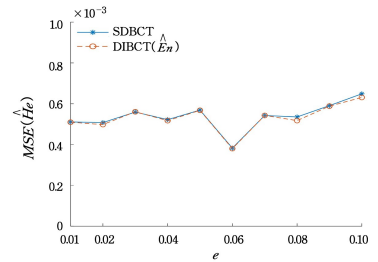


图 15  $e$  动态增加时估计值  $\hat{En}$  的均方差 ( $He=0.3$ )

Fig. 15  $MSE(\hat{En})$  for  $e$  dynamically increasing ( $He=0.3$ )

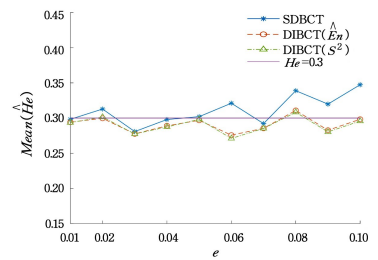


图 16  $e$  动态增加时估计值  $\hat{He}$  的均值 ( $He=0.3$ )

Fig. 16  $Mean(\hat{He})$  for  $e$  dynamically increasing ( $He=0.3$ )

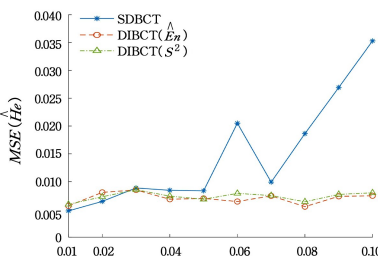


图 17  $e$  动态增加时估计值  $\hat{He}$  的均方差 ( $He=0.3$ )

Fig. 17  $MSE(\hat{He})$  for  $e$  dynamically increasing ( $He=0.3$ )

针对超熵  $He$  的估计情况,随着  $e$  增大,当  $He=0.1$  时,从图 12 可以看出 SDBCT 对超熵估计值的均值  $Mean(\hat{He})$  越来越大大于真值 0.1,而  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  对超熵估计值的均值  $Mean(\hat{He})$  在真值附近,没有增大趋势。进一步,从图 13 可以看出,SDBCT 的均方误差  $MSE(\hat{He})$  逐渐增大,速度较快,而  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  的  $MSE(\hat{He})$  并不因  $e$  的增大而发生较大变化。同样地,当  $He=0.3$  时,从图 16 和图 17 可以看出,各算法均值  $Mean(\hat{He})$  和均方误差  $MSE(\hat{He})$  有相似变化。这说明在估计  $He$  时,SDBCT 对参数  $e$  的变化比较敏感,而  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  对参数  $e$  的变化有较好的适应性。

#### 4.5 超熵 $He$ 和参数 $e$ 同时变化时算法的性能分析

首先,令  $(Ex=20, En=3, He)$ ,其中超熵  $He$  取值以 0.05 为步长从 0 增至 1.25。对不同超熵  $He$ ,利用算法 FCT 产生容量为  $N=10000$  的随机样本。其次,分别利用 SDBCT,  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  对熵  $En$  和超熵  $He$  进行估计。3 种算法中参数  $e$  都是以 0.01 为步长从 0.01 增加至 0.1。最后,将上述实验过程重复  $T=100$  次,分别计算各算法在  $T=100$  次的均方误差  $MSE(\hat{En})$  与  $MSE(\hat{He})$ ,具体结果分别如图 18 和图 19 所示。

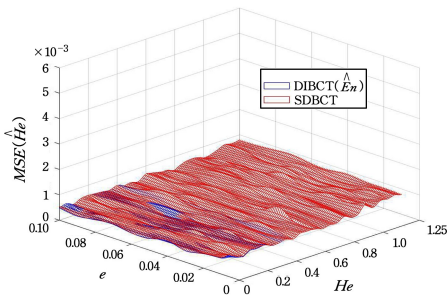


图 18  $He$  和  $e$  动态增加时估计值  $\hat{En}$  的均方误差

Fig. 18  $MSE(\hat{En})$  when  $He$  and  $e$  dynamically increase

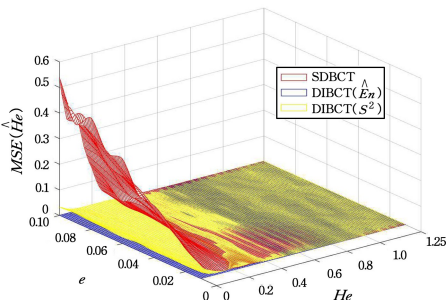


图 19  $He$  和  $e$  动态增加时估计值  $\hat{He}$  的均方误差(步长为 0.05)

Fig. 19  $MSE(\hat{He})$  when  $He$  and  $e$  dynamically increase ( $step=0.05$ )

从图 18 可以看出,超熵  $He$  和参数  $e$  同时增加时,SDBCT 和  $DIBCT(\hat{En})$  得到的均方误差  $MSE(\hat{En})$  波动都较小,说明这两种算法对熵  $En$  的估计值都比较稳定。从图 19 可以看出,当  $He$  和  $e$  增加时,算法 SDBCT 得到的均方误差

$MSE(\hat{He})$  波动较大。尤其当  $He$  处于  $[0, 0.5]$  时,随着  $e$  增大,SDBCT 得到的均方误差  $MSE(\hat{He})$  逐渐增大。相比之下, $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  得到的均方误差  $MSE(\hat{He})$  变化比较平稳。在  $He \geq 0.5$  时 3 种算法得到的均方误差  $MSE(\hat{He})$  趋于重合状态。

为进一步对比 SDBCT,  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  在参数  $e$  和超熵  $He$  同时变化时的效果,将超熵  $He$  的取值改为以 0.01 为步长从 0 增加至 0.5,其余条件不变,再次重复以上实验,结果如图 20 所示。

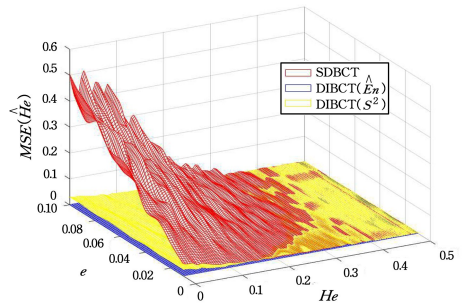


图 20  $He$  和  $e$  动态增加时估计值  $\hat{He}$  的均方误差(步长为 0.01)

Fig. 20  $MSE(\hat{He})$  when  $He$  and  $e$  dynamically increase( $step=0.01$ )

从图 20 可以看出,在参数  $e$  从 0.01 增加至 0.1 的过程中,SDBCT 得到的均方误差  $MSE(\hat{He})$  逐渐增大,而  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  得到的均方误差  $MSE(\hat{He})$  较小且稳定。这说明 SDBCT 对于  $He$  和  $e$  的变化较敏感,而  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  对于  $He$  和  $e$  的变化具有较强的鲁棒性。

## 5 应用举例

知识是人类认识客观世界实践经验的概况和总结。逆向云变换就是从细粒度数据中获取粗粒度知识的一种方式,是知识获取最基础的步骤,也是信息粒化的重要工具。因此,逆向云变换算法在综合评价、系统评估、可视化分析、图像分割、不确定性决策和聚类分析等领域得到了广泛应用。本章利用算法  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  对射击选手的射击水平进行还原和评价。以靶心为原点建立平面坐标系,将 A 和 B 两位选手每次射击的成绩看作一个二维云滴。A 和 B 两位选手分别进行 20 次射击,将这 20 次射击成绩看作 20 个二维云滴组成的云团,射击结果如图 21 所示。可以看出,选手 A 的射击结果向左下偏离,而且每次射击结果比较集中,结果相对稳定;而选手 B 的射击结果在靶心周围都有出现,且每次射击不够集中,整体不够稳定。为进一步观察 A 和 B 两位射手的射击水平,利用  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  分别估计选手 A 和选手 B 这 20 次射击成绩的数字特征  $(Ex, En, He)$  取值,然后将 A 和 B 两位选手射击水平的数字特征作为输入,利用正向云变换算法分别还原选手 A 和 B 20 次、100 次的随机射击成绩,结果分别如图 22 和图 23 所示。可以看出,还原成绩与实际成绩非常相近,说明  $DIBCT(\hat{En})$  和  $DIBCT(S^2)$  对于选手射击水平的评价是可信的。

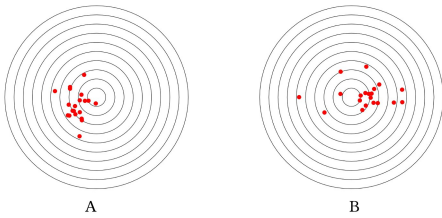
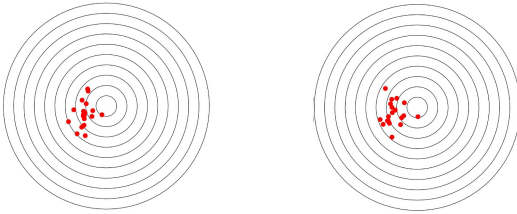


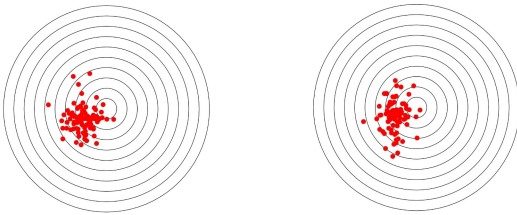
图 21 选手 A 和 B 的 20 次射击成绩

Fig. 21 20 shooting results of player A and B



(a)  $\hat{DIBCT}(En)$  还原选手 A  
20 次射击成绩

(b)  $\hat{DIBCT}(S^2)$  还原选手 A  
20 次射击成绩

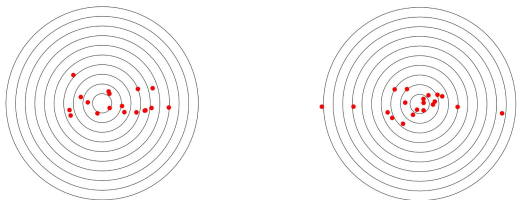


(c)  $\hat{DIBCT}(En)$  还原选手 A  
100 次射击成绩

(d)  $\hat{DIBCT}(S^2)$  还原选手 A  
100 次射击成绩

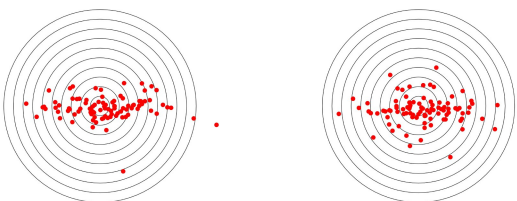
图 22 选手 A 射击成绩的还原

Fig. 22 Restore the score of player A



(a)  $\hat{DIBCT}(En)$  还原选手 B  
20 次射击成绩

(b)  $\hat{DIBCT}(S^2)$  还原选手 B  
20 次射击成绩



(c)  $\hat{DIBCT}(En)$  还原选手 B  
100 次射击成绩

(d)  $\hat{DIBCT}(S^2)$  还原选手 B  
100 次射击成绩

图 23 选手 B 射击成绩的还原

Fig. 23 Restore the score of player B

**结束语** 本文首先对经典的一阶绝对中心矩逆向云变换算法 SBCT-1stM 的不足进行实验和理论分析,发现超熵估计值为虚数的原因是由熵估计方法的不足导致的。其次,受正向云变换算法生成云滴特点和算法 SDBCT 思想的启发,将正态随机变量动态产生新样本融合到原始样本数据中,从而

提出了修正熵  $En$  和修正方差  $S^2$  的动态增量式逆向云变换算法  $\hat{DIBCT}(En)$  和  $\hat{DIBCT}(S^2)$ 。利用随机模拟实验将提出的两种算法  $\hat{DIBCT}(En)$  和  $\hat{DIBCT}(S^2)$  与已有算法进行了对比分析,实验结果表明算法  $\hat{DIBCT}(En)$  和  $\hat{DIBCT}(S^2)$  对参数估计的误差较小,且具有较好的稳定性、收敛性和对参数的鲁棒性。最后,将提出的两种算法用于对射击选手射击水平的评价中,结果表明了算法  $\hat{DIBCT}(En)$  和  $\hat{DIBCT}(S^2)$  的实用性。但是,本文仅从实验模拟分析角度对所提算法进行了模拟分析,增量参数  $e$  对算法的影响机制还有待进一步探究和分析。同时,由于二阶正态云一阶绝对中心矩存在不解析的情况,因此从数值分析的角度对逆向云变换算法进行研究是一个研究方向。此外,结合算法  $\hat{DIBCT}(En)$  和  $\hat{DIBCT}(S^2)$  的特点,将其应用于增量数据挖掘和动态增量三支决策都是下一步需要研究的重点。

### 参 考 文 献

- [1] LI D Y. Ten questions and answers for the new generation of artificial intelligences [J]. CAAI Transactions on Intelligent Systems, 2021, 16(5): 828-833.
- [2] WANG G Y, FU S, YANG J, et al. A review of research on multi-granularity based intelligent computing [J]. Chinese Journal of Computers, 2022, 45(6): 1161-1175.
- [3] QIN J D, MARTINEZ L, PEDRYCZ W, et al. An overview of granular computing in decision-making: Extensions, applications, and challenges [J]. Information Fusion, 2023, 98: 101833.
- [4] AGGARWAL M. An entropy framework for randomness and fuzziness [J]. Expert Systems with Applications, 2024, 243: 122431.
- [5] TAGHIKHANI S, BAROUGHI F. Fuzzy random classical and inverse median location problems [J]. Soft Computing, 2023, 27: 8821-8839.
- [6] ZADEH L. Fuzzy sets [J]. Information and Control, 1965, 8(3): 338-353.
- [7] PAWLAK Z. Rough sets [J]. International Journal of Computer and Information Sciences, 1982, 11(5): 341-356.
- [8] YANG X, LI Y, LI T. A review of sequential three-way decision and multi-granularity learning [J]. International Journal of Approximate Reasoning, 2023, 152: 414-433.
- [9] LI D Y, LIU C Y, GAN W Y. Proof of the heavy-tailed property of normal cloud model [J]. Strategic Study of CAE, 2011, 13(4): 20-23.
- [10] CHAUHAN P, GUPTA A, MALHOTRA T. A novel cloud model based on multiplicative unbalanced linguistic term set [J]. The Journal of Supercomputing, 2023, 79(14): 16378-16408.
- [11] GAO H B, ZHANG X Y, ZHANG T L, et al. Research of intelligent vehicle variable granularity evaluation based on cloud model [J]. Acta Electronica Sinica, 2016, 44(2): 365-373.
- [12] DAI J, HU B, WANG G Y, et al. The uncertainty similarity measure of cloud model based on the fusion of distribution contour and local feature [J]. Journal of Electronics & Information

- Technology, 2022, 44(4):1429-1439.
- [13] JIANG J, YUAN Y P, DENG P, et al. Effectiveness evaluation of Chongqing shipping center project with cloud model and VIKOR method [J]. Navigation of China, 2023, 46(3): 91-97, 104.
- [14] MA W, LI Y R. Analysis of a performance evaluation method for science and technology investment projects using SAHP cloud model [J]. Journal of Southwest University(Natural Science Edition), 2023, 45(10):160-168.
- [15] LIU J K, WANG J R, WANG C X. Research on the subject and method of emergency capability evaluation based on bibliometric analysis [J]. Journal of Safety and Environment, 2023, 23(5): 1398-1406.
- [16] XU C L, XU H. Research on similarity measurement method of normal cloud based on Hellinger distance and its application[J]. CAAI Transactions on Intelligent Systems, 2023, 18(6): 1312-1321.
- [17] LIU W Q, ZHU J J. A multistage decision-making method for multi-source information with Shapley optimization based on normal cloud models[J]. Applied Soft Computing, 2021, 111: 107716.
- [18] LIU W Q, ZHU J J. A multistage decision-making method with quantum-guided expert state transition based on normal cloud models[J]. Information Sciences, 2022, 615:700-730.
- [19] ZHANG L M, CHEN W Y. Multi-criteria group decision-making with cloud model and TOPSIS for alternative selection under uncertainty[J]. Soft Computing, 2022, 26:12509-12529.
- [20] LIU Z M, WANG X Y, WANG W X, et al. An integrated TOPSIS-ORESTE-based decision-making framework for new energy investment assessment with cloud model[J]. Computational and Applied Mathematics, 2022, 41:42.
- [21] LIU Y, LIU Z T, LI S, et al. Cloud-Cluster: An uncertainty clustering algorithm based on cloud model [J]. Knowledge-Based Systems, 2023, 263:110261.
- [22] LIU W Q, ZHU J J, CHICLANA F. Large-scale group consensus hybrid strategies with three-dimensional clustering optimisation based on normal cloud models [J]. Information Fusion, 2023, 94:66-91.
- [23] XU C L, WANG G Y. Bidirectional cognitive computing model for uncertain concepts[J]. Cognitive Computation, 2019, 11(5): 613-629.
- [24] LIU C Y, FENG M, DAI X J, et al. A new algorithm of backward cloud [J]. Journal of System Simulation, 2004(11):2417-2420.
- [25] YU S W, SHI Z K. New algorithm of backward cloud based on normal interval number [J]. Systems Engineering-Theory & Practice, 2011, 31(10):2021-2026.
- [26] XU C L, WANG G Y, ZHANG Q H. A new multi-step backward cloud transformation algorithm based on normal cloud model[J]. Fundamenta Informaticae, 2014, 133:55-85.
- [27] XU C L, WANG G Y. A novel cognitive transformation algorithm based on Gaussian cloud model and its application in image segmentation [J]. Numerical Algorithms, 2017, 76(4): 1039-1070.
- [28] CHEN H, LI B, LIU C Y. An algorithm of backward cloud without certainty degree [J]. Journal of Chinese Computer Systems, 2015, 36(3):544-549.
- [29] LI D Y, LIU C Y, GAN W Y. A new cognitive model: cloud model[J]. International Journal of Intelligent Systems, 2009, 24(3):357-375.
- [30] YANG J, WANG G Y, LIU Q, et al. Retrospect and prospect of research of normal cloud model[J]. Chinese Journal of Computers, 2018, 41(3):724-744.
- [31] XU C L. Research on bidirectional cognitive computing method based on Cloud Model [D]. Chengdu: Southwest Jiaotong University, 2014.
- [32] GAO Y P, LU W Y, WANG L L, et al. Parameter estimation of cloud model based on Bayesian theory [J]. Statistics & Decision, 2019, 35(6):5-8.



**XU Changlin**, born in 1983, Ph.D, associate professor, master's supervisor, is a member of CCF (No. R4034M). His main research interests include intelligent information processing, cloud model theory, cognitive computing and uncertainty decision-making.

(责任编辑:何杨)