

函数P-集合与内P-信息规律依赖内-挖掘

赵树理^{1,2} 张环理¹ 史开泉¹

(山东大学数学与系统科学学院 济南 250100)¹ (商丘师范学院数学与信息科学学院 商丘 476000)²

摘要 函数P-集合(function packet sets)是把函数概念引入到P-集合(packet sets)内,改进P-集合得到的,函数P-集合具有动态特性、规律(函数)特性。函数P-集合是由函数内P-集合 S^F (function internal packet set S^F)与函数外P-集合 S^F (function outer packet set S^F)构成的函数集合对;或者, (S^F, S^F) 是函数P-集合。利用函数内P-集合给出内P-信息规律生成概念、内P-信息规律生成的属性定理、内P-信息规律依赖特征与数据损失定理、内P-信息规律依赖生成的信息规律变轨特征,并给出未知内P-信息规律依赖内-挖掘及其应用。

关键词 函数P-集合,内P-信息规律,规律依赖,依赖内-挖掘,信息规律变轨,应用

中图分类号 O144 **文献标识码** A

Function P-sets and Internal-mining of Internal P-information Law Dependence

ZHAO Shu-li^{1,2} ZHANG Huan-li¹ SHI Kai-quan¹

(School of Mathematics and System Sciences, Shandong University, Jinan 250100, China)¹

(School of Mathematics and Information Sciences, Shangqiu Normal University, Shangqiu 476000, China)²

Abstract Function P-sets(function packet sets) are obtained by introducing the function concept into and improving P-sets(packet sets), which have the dynamic characteristic. Function P-sets are a set-pair composed of function internal P-set S^F (function internal packet set S^F) and function outer P-set S^F (function outer packet set S^F), or (S^F, S^F) is function P-sets. Using function internal P-sets, the paper gave the concept and attribute theorem, the dependence characteristic and data loss theorem of internal P-information law generation, the orbital transfer characteristic of information law of internal P-information law dependence. At last, the internal-mining of unknown internal P-information law dependence and its application were given.

Keywords Function P-sets, Internal P-information law, Law dependence, Internal-mining, Orbital transfer of information law, Application

1 引言

函数P-集合(function packet sets)^[1,2]是用有限普通函数集合S代替P-集合^[4,5]中的有限普通元素集合X,把函数概念引入到P-集合内,改进P-集合得到的,函数P-集合具有动态特性与规律(函数)特性,在一类动态信息规律研究中得到了应用^[1-3];P-集合在一类动态信息研究中得到了应用^[4-22]。函数P-集合是由函数内P-集合 S^F (function internal packet set S^F)与函数外P-集合 S^F (function outer packet set S^F)构成的函数集合对;或者, (S^F, S^F) 是函数P-集合。在一定的条件下,函数P-集合被还原成有限普通函数集合S。函数P-集合具有动态特性,规律特性:给定有限普通函数集合S, α 是S的属性集合;1°若在 α 内补充属性, α 变成 $\alpha_1^F \subseteq \alpha_2^F \subseteq \dots \subseteq \alpha_{n-1}^F \subseteq \alpha_n^F$;则S变成函数内P-集合 $S_n^F \subseteq S_{n-1}^F \subseteq \dots \subseteq S_2^F \subseteq S_1^F$;2°若在 α 内删除属性, α 变成 $\alpha_n^F \subseteq \alpha_{n-1}^F \subseteq \dots \subseteq \alpha_2^F \subseteq \alpha_1^F$;则S变成函数外P-集合 $S_1^F \subseteq S_2^F \subseteq \dots \subseteq S_{n-1}^F \subseteq S_n^F$ 。把单依赖“ $\Rightarrow$ ”概念引

入到1°与2°中,由1°得到:若 $\alpha_1^F \Rightarrow \alpha_2^F \Rightarrow \dots \Rightarrow \alpha_{n-1}^F \Rightarrow \alpha_n^F$,则 $S_n^F \Rightarrow S_{n-1}^F \Rightarrow \dots \Rightarrow S_2^F \Rightarrow S_1^F$;由2°得到:若 $\alpha_n^F \Rightarrow \alpha_{n-1}^F \Rightarrow \dots \Rightarrow \alpha_2^F \Rightarrow \alpha_1^F$,则 $S_1^F \Rightarrow S_2^F \Rightarrow \dots \Rightarrow S_{n-1}^F \Rightarrow S_n^F$;或者, S_{k-1}^F 单依赖于 S_k^F , S_k^F 单依赖于 S_{k-1}^F 。如果把函数内P-集合 S^F 定义成内P-信息规律 $w(x)^F$,函数外P-集合 S^F 定义成外P-信息规律 $w(x)^F$;则 $w(x)_{k-1}^F$ 单依赖于 $w(x)_k^F$, $w(x)_k^F \Rightarrow w(x)_{k-1}^F$; $w(x)_k^F$ 单依赖于 $w(x)_{k-1}^F$, $w(x)_{k-1}^F \Rightarrow w(x)_k^F$ 。“信息规律依赖特性”潜藏在函数P-集合内,未引起人们的重视,人们也未对它进行研究,给出它的应用。

本文给出这样的研究:如果知道内P-信息规律 $w(x)_{k-1}^F$,利用信息规律依赖特性怎样找到(挖掘到)内P-信息规律 $w(x)_k^F$? $w(x)_k^F$ 是不被人们事先知道的。如果找到 $w(x)_k^F$, $w(x)_k^F$ 的属性集合 α_k^F 与 $w(x)_{k-1}^F$ 的属性集合 α_{k-1}^F 之间具有什么样的关系? 如果 $w(x)_{k-1}^F$ 单依赖于 $w(x)_k^F$;或者, $w(x)_k^F \Rightarrow w(x)_{k-1}^F$,则 $w(x)_k^F$ 的形状发生变化($w(x)_k^F$ 出现变轨), $w(x)_k^F$ 的变轨特征是什么? 具有什么样的应用? 利用函数内

到稿日期:2012-07-21 返修日期:2013-02-01 本文受福建省龙岩市科技计划基金项目(2011LY20),河南省基础与前沿技术研究项目(122300410423)资助。

赵树理(1962—),男,副教授,主要研究方向为信息系统与信息识别,E-mail: zhaolishuli@163.com;张环理(1963—),男,副教授,主要研究方向为信息系统与应用;史开泉(1945—),男,教授,博士生导师,主要研究方向为信息系统理论与应用,E-mail: shikq@sdu.edu.cn(通信作者)。

P-集合,给出这些问题的讨论,本文的主要结果:给出内 P-信息规律生成、内 P-信息规律生成的属性定理、内 P-信息规律依赖特征与数据损失定理、内 P-信息规律依赖生成的信息规律变轨特征,并给出未知内 P-信息规律依赖内-挖掘及其应用。

为了便于讨论,方便读者,把函数 P-集合简单地引入到本文的第 2 节内,作为本文讨论的预备知识,函数 P-集合的更多特征,见文献[1,2]。

2 函数 P-集合

约定 $U(x)$ 是有限函数论域; $V(\alpha)$ 是有限属性论域; $S(x) = \{s(x)_1, s(x)_2, \dots, s(x)_n\}$ 是 $U(x)$ 上的有限函数集合; $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ 是 $V(\alpha)$ 上的有限属性集合; $s(x), r(x)$ 是 x 的函数。为了简单, $U(x), V(\alpha), S(x), r(x)$ 分别记作 U, V, S, r 。

文献[1,2]给出:

给定函数集合 $S = \{s_1, s_2, \dots, s_q\} \subset U, \alpha = \{\alpha_1, \alpha_2, \dots, \alpha_k\} \subset V$ 是 S 的属性集合,称 S^F 是 S 生成的函数内 P-集合(function internal packet set),而且

$$S^F = S - S^- \quad (1)$$

S^- 称作 S 的 $\bar{F}$ -函数(元素)删除集合,而且

$$S^- = \{s_i | s_i \in S, \bar{f}(s_i) = r_i \in S, \bar{f} \in \bar{F}\} \quad (2)$$

如果 S^F 具有的属性集合 α^F 满足

$$\alpha^F = \alpha \cup \{\alpha_i' | f(\beta_i) = \alpha_i' \in \alpha, f \in F\} \quad (3)$$

这里:式(3)中, $\beta_i \in V, \beta_i \in \alpha, f \in F$ 把 β_i 变成 $f(\beta_i) = \alpha_i' \in \alpha, S^F \neq \emptyset$; 式(1)中, $S^F = \{s_1, s_2, \dots, s_p\}, p \leq q, p, q \in \mathbb{N}^+$ 。

给定函数集合 $S = \{s_1, s_2, \dots, s_q\} \subset U, \alpha = \{\alpha_1, \alpha_2, \dots, \alpha_k\} \subset V$ 是 S 的属性集合,称 S^F 是 S 生成的函数外 P-集合(function outer packet set),而且

$$S^F = S \cup S^+ \quad (4)$$

S^+ 称作 S 的 F -函数(元素)补充集合,而且

$$S^+ = \{r_i | r_i \in U, r_i \in S, f(r_i) = s_i' \in S, f \in F\} \quad (5)$$

如果 S^F 具有的属性集合 α^F 满足

$$\alpha^F = \alpha - \{\beta_i | \bar{f}(\alpha_i) = \beta_i \in \alpha, \bar{f} \in \bar{F}\} \quad (6)$$

这里:式(6)中, $\alpha_i \in \alpha, \bar{f} \in \bar{F}$ 把 α_i 变成 $\bar{f}(\alpha_i) = \beta_i \in \alpha, \alpha^F \neq \emptyset$; 式(4)中, $S^F = \{s_1, s_2, \dots, s_r\}, q \leq r, q, r \in \mathbb{N}^+$ 。

由 S^F 与 S^F 构成的函数集合对,称作函数集合 S 生成的函数 P-集合(function packet sets),而且

$$(S^F, S^F) \quad (7)$$

有限函数集合 S 称作函数 P-集合(S^F, S^F)的基集合(基础集合)。

利用式(3),在 α 内不断地补充属性,得到

$$\alpha_1^F \subseteq \alpha_2^F \subseteq \dots \subseteq \alpha_{n-1}^F \subseteq \alpha_n^F \quad (8)$$

由式(8)得到函数内 P-集合:

$$S_n^F \subseteq S_{n-1}^F \subseteq \dots \subseteq S_2^F \subseteq S_1^F \quad (9)$$

利用式(6),在 α 内不断地删除属性,得到:

$$\alpha_n^F \subseteq \alpha_{n-1}^F \subseteq \dots \subseteq \alpha_2^F \subseteq \alpha_1^F \quad (10)$$

由式(10)得到函数外 P-集合:

$$S_1^F \subseteq S_2^F \subseteq \dots \subseteq S_{n-1}^F \subseteq S_n^F \quad (11)$$

利用式(9),式(11)得到:

$$\{(S_i^F, S_j^F) | i \in I, j \in J\} \quad (12)$$

称 $\{(S_i^F, S_j^F) | i \in I, j \in J\}$ 是函数 P-集合的集合对族,它是函

数 P-集合的一般表达式; I, J 是指标集合。

利用式(1)~式(12),容易证明:

定理 1(函数 P-集合第一还原定理) 若 $F = \bar{F} = \emptyset$, 则函数 P-集合被还原成有限函数集合 S ; 或者,

$$(S^F, S^F)_{F=\bar{F}=\emptyset} = S \quad (13)$$

定理 2(函数 P-集合第二还原定理) 若 $F = \bar{F} = \emptyset$, 则函数 P-集合被还原成有限函数集合 S ; 或者,

$$\{(S_i^F, S_j^F) | i \in I, j \in J\}_{F=\bar{F}=\emptyset} = S \quad (14)$$

命题 1 在动态-静态条件下,函数 P-集合是有限函数集合 S 的一般形式,有限函数集合 S 是函数 P-集合的特例。

命题 2 函数 P-集合的存在依赖于函数集合 S 的属性集合 α 内的属性补充与属性删除。

应当特别指出:式(2)、式(3)、式(5)、式(6)中的 $F = \{f_1, f_2, \dots, f_m\}, \bar{F} = \{\bar{f}_1, \bar{f}_2, \dots, \bar{f}_n\}$ 是函数(属性迁移)族; $f \in F, \bar{f} \in \bar{F}$ 是函数(属性)迁移。 $f \in F$ 的特征:对于函数 $r_i \in U, r_i \in S, f(r_i) = s_i' \in S$; 对于属性 $\beta_i \in V, \beta_i \in \alpha, f(\beta_i) = \alpha_i' \in \alpha$ 。 $\bar{f} \in \bar{F}$ 的特征是:对于函数 $s_i \in S, \bar{f}(s_i) = r_i \in S$; 对于属性 $\alpha_i \in \alpha, \bar{f}(\alpha_i) = \beta_i \in \alpha$ 。 $f \in F, \bar{f} \in \bar{F}$ 更多的概念与意义,见文献[1,2]。

利用第 2 节中的概念,第 3 节中给出如下。

3 P-信息规律生成与它的属性补充-删除特征

定义 1 给定有限普通函数集合 $S = \{s_1, s_2, \dots, s_q\}, \forall y_i = \{y_{i,1}, y_{i,2}, \dots, y_{i,n}\}$ 是 $s_i \in S$ 生成的离散数据集合,称 y 是 S 生成的离散数据集合,而且

$$y = \{y_1, y_2, \dots, y_n\} = \{\sum_{i=1}^q y_{i,1}, \sum_{i=1}^q y_{i,2}, \dots, \sum_{i=1}^q y_{i,n}\} \quad (15)$$

定义 2 称

$$\begin{aligned} w(x) &= \sum_{j=1}^n y_j \prod_{\substack{i=1 \\ i \neq j}}^n \frac{x - x_i}{x_j - x_i} \\ &= a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + a_1 x + a_0 \end{aligned} \quad (16)$$

是 S 生成的信息规律,如果 $w(x)$ 是通过 y 生成的数据点 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 插值得到的。

定义 3 $w(x)^F, w(x)^{\bar{F}}$ 分别称作函数内 P-集合 S^F 、函数外 P-集合 S^F 生成的内 P-信息规律、外 P-信息规律,而且

$$w(x)^F = b_{n-1} x^{n-1} + b_{n-2} x^{n-2} + \dots + b_1 x + b_0 \quad (17)$$

$$w(x)^{\bar{F}} = c_{n-1} x^{n-1} + c_{n-2} x^{n-2} + \dots + c_1 x + c_0 \quad (18)$$

这里: $w(x)^F$ 是通过数据点 $(x_1, y_1^F), (x_2, y_2^F), \dots, (x_n, y_n^F)$ 由式(16)得到的; $(x_1, y_1^F), (x_2, y_2^F), \dots, (x_n, y_n^F)$ 是由离散数据集合 $y^F = \{y_1^F, y_2^F, \dots, y_n^F\} = \{\sum_{i=1}^q y_{i,1}, \sum_{i=1}^q y_{i,2}, \dots, \sum_{i=1}^q y_{i,n}\}$ 生成的数据点, y^F 是 S^F 生成的离散数据集合。

$w(x)^{\bar{F}}$ 是通过数据点 $(x_1, y_1^{\bar{F}}), (x_2, y_2^{\bar{F}}), \dots, (x_n, y_n^{\bar{F}})$ 由式(16)得到的; $(x_1, y_1^{\bar{F}}), (x_2, y_2^{\bar{F}}), \dots, (x_n, y_n^{\bar{F}})$ 是由离散数据集合

$y^{\bar{F}} = \{y_1^{\bar{F}}, y_2^{\bar{F}}, \dots, y_n^{\bar{F}}\} = \{\sum_{i=1}^q y_{i,1}, \sum_{i=1}^q y_{i,2}, \dots, \sum_{i=1}^q y_{i,n}\}$ 生成的数据点, $y^{\bar{F}}$ 是 S^F 生成的离散数据集合, $p \leq q \leq r, p, q, r \in \mathbb{N}^+$ 。

定义 4 由内 P-信息规律 $w(x)^F$ 与外 P-信息规律 $w(x)^{\bar{F}}$ 构成的信息规律对,称作函数 P-集合生成的 P-信息规律,而且

$$(w(x)^F, w(x)^{\bar{F}}) \quad (19)$$

定义 5 称

$$\{(w(x)_i^F, w(x)_j^{\bar{F}}) | i \in I, j \in J\} \quad (20)$$

是函数 P-集合生成的 P-信息规律族。

由式(15)一式(20)得到:

命题 3 内 P-信息规律 $w(x)^F$ 与信息规律 $w(x)$ 满足 $w(x)^F - w(x) \leq 0$ 。

命题 4 外 P-信息规律 $w(x)^F$ 与信息规律 $w(x)$ 满足 $w(x)^F - w(x) \geq 0$ 。

命题 5 内 P-信息规律 $w(x)^F$ 、信息规律 $w(x)$ 与外 P-信息规律 $w(x)^F$ 满足 $w(x)^F \leq w(x) \leq w(x)^F$ 。

定理 3(内 P-信息规律生成与属性补充定理) $w(x)^F$ 是 S^F 生成的内 P-信息规律的充分必要条件是存在属性集合 $\Delta\alpha \neq \emptyset$, S 的属性集合 α , S^F 的属性集合 α^F 与 $\Delta\alpha$ 满足

$$\alpha^F = \alpha \cup \Delta\alpha \quad (21)$$

证明: $w(x)^F$ 是 S^F 生成的内 P-信息规律, $w(x)^F$ 与 S^F 具有相同的属性集合 α^F ; $w(x)$ 是 S 生成的信息规律, $w(x)$ 与 S 具有相同的属性集合 α 。1°由式(1)一式(3), 式(16), 式(17)得到: S^F 的属性集合 α^F 与 S 的属性集合 α 满足 $\alpha \subseteq \alpha^F$; 或者 $w(x)^F$ 的属性集合 α^F 与 $w(x)$ 的属性集合 α 满足 $\alpha \subseteq \alpha^F$; 显然存在属性集合 $\Delta\alpha \neq \emptyset$, $\Delta\alpha = \alpha^F - \alpha$; 或者, $\alpha^F = \alpha \cup \Delta\alpha$ 。2°给定属性集合 $\alpha, \alpha^F, \alpha \subseteq \alpha^F$; 显然存在属性集合 $\Delta\alpha \neq \emptyset, \alpha \cup \Delta\alpha = \alpha^F$; 若 α, α^F 分别是 $w(x), w(x)^F$ 的属性集合, 由式(1)一式(3), 式(16), 式(17)得到: $w(x)^F$ 是 S^F 生成的内 P-信息规律, $w(x)$ 是 S 生成的信息规律; 由 1°与 2°, 得到定理 3。

定理 4(外 P-信息规律生成与属性删除定理) $w(x)^F$ 是 S^F 生成的外 P-信息规律的充分必要条件是存在属性集合 $\nabla\alpha \neq \emptyset$, S 的属性集合 α , S^F 的属性集合 α^F 与 $\nabla\alpha$ 满足

$$\alpha^F = \alpha - \nabla\alpha \quad (22)$$

定理 4 的证明与定理 3 类似, 证明略。

由定理 3、定理 4 直接得到:

定理 5(P-信息规律生成与属性补充-删除定理) 若 $(w(x)^F, w(x)^F)$ 是 (S^F, S^F) 生成的 P-信息规律, 则存在属性集合 $(\Delta\alpha, \nabla\alpha) \neq \emptyset$, (S^F, S^F) 的属性集合 (α^F, α^F) 与 $(\Delta\alpha, \nabla\alpha)$ 满足

$$(\alpha^F - \Delta\alpha) - (\alpha^F \cup \nabla\alpha) = \emptyset \quad (23)$$

这里: 式(23)表示: $\alpha^F - \Delta\alpha \neq \emptyset$; $(\Delta\alpha, \nabla\alpha) \neq \emptyset$ 表示: $\Delta\alpha \neq \emptyset$, $\nabla\alpha \neq \emptyset$ 。

由定理 3—定理 5 分别得到:

推论 1 若 $\{w(x)_i^F | i \in I\}$ 是 S^F 生成的内 P-信息规律族, 则

$$w(x)_\lambda^F = \min_{i \in I} (w(x)_i^F) \quad (24)$$

的属性集合 α_λ^F 满足

$$\alpha_\lambda^F = \max_{i \in I} (\alpha_i^F) \quad (25)$$

推论 2 若 $\{w(x)_j^F | j \in J\}$ 是 S^F 生成的外 P-信息规律族, 则

$$w(x)_\tau^F = \max_{j \in J} (w(x)_j^F) \quad (26)$$

的属性集合 α_τ^F 满足

$$\alpha_\tau^F = \min_{j \in J} (\alpha_j^F) \quad (27)$$

推论 3 $\{(w(x)_i^F, w(x)_j^F) | i \in I, j \in J\}$ 中, 存在 $w(x)_\lambda^F$, $w(x)_\tau^F$; $w(x)_\lambda^F$ 具有最大属性集合 α_λ^F , $w(x)_\tau^F$ 具有最小属性集合 α_τ^F 。

图 1 给出内 P-信息规律 $w(x)_i^F$ 、外 P-信息规律 $w(x)_j^F$ 、信息规律 $w(x)$ 的直观表示; $w(x)_i^F, w(x)_j^F, w(x)$ 用折线表示。

利用第 2 节, 第 3 节中的概念, 第 4 节—第 5 节中给出内 P-信息规律依赖、内 P-信息规律依赖变轨与未知内 P-信息规律依赖挖掘的应用。

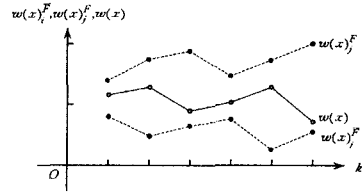


图 1
 $w(x)$ 是 S 生成的信息规律, $w(x)$ 具有属性集合 α , $w(x)$ 用实线表示;
 $w(x)_i^F$ 是 S_i^F 生成的内 P-信息规律, $w(x)_i^F$ 具有属性集合 $\alpha_i^F, \alpha \subseteq \alpha_i^F$;
 $w(x)_j^F$ 是 S_j^F 生成的外 P-信息规律, $w(x)_j^F$ 具有属性集合 $\alpha_j^F, \alpha_j^F \subseteq \alpha$;
 $w(x)_i^F, w(x)_j^F$ 用虚线表示。

图 1

4 内 P-信息规律依赖与依赖性定理

定义 6 称内 P-信息规律 $w(x)_k^F$ 单依赖于 $w(x)_{k+1}^F$, 记作

$$w(x)_{k+1}^F \Rightarrow w(x)_k^F \quad (28)$$

如果生成 $w(x)_k^F, w(x)_{k+1}^F$ 的函数内 P-集合 S_k^F, S_{k+1}^F 满足

$$S_{k+1}^F \Rightarrow S_k^F \quad (29)$$

定义 7 称内 P-信息规律 $w(x)_k^F$ 双依赖于 $w(x)_{k+1}^F$, 记作

$$w(x)_{k+1}^F \Leftrightarrow w(x)_k^F \quad (30)$$

如果生成 $w(x)_k^F, w(x)_{k+1}^F$ 的函数内 P-集合 S_k^F, S_{k+1}^F 满足

$$S_{k+1}^F \Leftrightarrow S_k^F \quad (31)$$

这里: 式(29), 式(31)中, “ $\Rightarrow$ ”与“ $\subseteq$ ”等价; “ $\Leftrightarrow$ ”与“ $=$ ”等价; “单依赖 $\Rightarrow$ ”, “双依赖 $\Leftrightarrow$ ”取自数理逻辑。为了简化, 符号的统一, 又不引起误解和混乱, 式(28), 式(30)中分别取与式(29), 式(31)中相同的记号; 在式(28)中, “ $\Rightarrow$ ”与“ $\leq$ ”等价, 式(30)中, “ $\Leftrightarrow$ ”与“ $=$ ”等价。

定义 8 称 $\nabla w(x)_k^F \neq 0$ 是 $w(x)_k^F$ 单依赖于 $w(x)_{k+1}^F$ 生成的单依赖损失, 如果 $w(x)_{k+1}^F$ 与 $w(x)_k^F$ 满足

$$w(x)_k^F - w(x)_{k+1}^F = \nabla w(x)_k^F \quad (32)$$

由定义 6—定义 8 得到:

定理 6(内 P-信息规律单依赖辨识定理) 若内 P-信息规律 $w(x)_k^F$ 单依赖于 $w(x)_{k+1}^F$, 则

$$\text{IDE}(w(x)_{k+1}^F, w(x)_k^F) \quad (33)$$

定理 7(内 P-信息规律双依赖不可辨识定理) 若内 P-信息规律 $w(x)_k^F$ 双依赖于 $w(x)_{k+1}^F$, 则

$$\text{UNI}(w(x)_{k+1}^F, w(x)_k^F) \quad (34)$$

这里: 式(33), 式(34)中, IDE=identification, UNI=unidentification。

定理 8(内 P-信息规律单依赖-双依赖定理) 若内 P-信息规律 $w(x)_k^F$ 与 $w(x)_{k+1}^F$ 满足

$$w(x)_{k+1}^F \Rightarrow w(x)_k^F \quad (35)$$

则存在单依赖损失 $\nabla w(x)_k^F \neq 0$, $(w(x)_k^F - \nabla w(x)_k^F)$ 双依赖于 $w(x)_{k+1}^F$; 或者

$$w(x)_{k+1}^F \Leftrightarrow (w(x)_k^F - \nabla w(x)_k^F) \quad (36)$$

定理 6—定理 8 是直接的结果, 它们的证明, 略。

定理 9(内 P-信息规律单依赖数据损失定理) 内 P-信息规律 $w(x)_k^F$ 单依赖于 $w(x)_{k+1}^F$ 的充分必要条件是存在数据

集合 $\nabla y_k^F; \nabla y_k^F, y_k^F$ 与 y_{k+1}^F 满足

$$\nabla y_k^F = y_k^F - y_{k+1}^F \quad (37)$$

式中, ∇y_k^F 是 $w(x)_k^F$ 单依赖于 $w(x)_{k+1}^F$ 生成的数据损失集合。

这里:式(37)中, $\nabla y_k^F = y_k^F - y_{k+1}^F = \{y_{k,1}^F - y_{k+1,1}^F, y_{k,2}^F - y_{k+1,2}^F, \dots, y_{k,n}^F - y_{k+1,n}^F\}, \forall r \in (1, 2, \dots, n), y_{k,r}^F - y_{k+1,r}^F \neq 0; y_k^F, y_{k+1}^F$ 分别是生成 $w(x)_k^F, w(x)_{k+1}^F$ 的数据集合。

证明:1°给定内P-信息规律 $w(x)_k^F, w(x)_{k+1}^F; w(x)_k^F$ 单依赖于 $w(x)_{k+1}^F$;或者, $w(x)_{k+1}^F \Rightarrow w(x)_k^F$;由图1直接得到: $w(x)_{k+1}^F \leq w(x)_k^F$ 。对 $w(x)_k^F, w(x)_{k+1}^F$ 离散得到数据集合: $y_k^F = \{y_{k,1}^F, y_{k,2}^F, \dots, y_{k,n}^F\}, y_{k+1}^F = \{y_{k+1,1}^F, y_{k+1,2}^F, \dots, y_{k+1,n}^F\}$,则有 $\forall p \in (1, 2, \dots, n), y_{k+1,p}^F \leq y_{k,p}^F$;存在 $\nabla y_{k,p}^F \neq 0, y_{k,p}^F - y_{k+1,p}^F = \nabla y_{k,p}^F$;或者, $\nabla y_k^F = y_k^F - y_{k+1}^F$ 。2°给定数据集合 $y_k^F, y_{k+1}^F, \nabla y_k^F \neq 0$,满足 $\nabla y_k^F = y_k^F - y_{k+1}^F$;或者, $y_{k+1}^F \leq y_k^F$;或者, $\forall y_{k+1,p}^F \leq y_{k,p}^F, y_{k+1,p}^F \in y_{k+1}^F, y_{k,p}^F \in y_k^F, p \in (1, 2, \dots, n)$;利用式(16),式(17)与 y_{k+1}^F, y_k^F 分别得到: $w(x)_{k+1}^F, w(x)_k^F$; $w(x)_{k+1}^F, w(x)_k^F$ 分别是 y_{k+1}^F, y_k^F 生成的内P-信息规律,而且 $w(x)_{k+1}^F \leq w(x)_k^F$;或者, $w(x)_{k+1}^F \Rightarrow w(x)_k^F$; $w(x)_k^F$ 单依赖于 $w(x)_{k+1}^F$;由1°与2°得到定理9。

利用式(28)一式(37)与式(21)容易得到:

推论4 若 $w(x)_k^F$ 单依赖于 $w(x)_{k+1}^F, w(x)_{k+1}^F \Rightarrow w(x)_k^F$,则存在属性集合 $\Delta \alpha_k^F \neq \emptyset, w(x)_k^F$ 的属性集合 $\alpha_k^F, w(x)_{k+1}^F$ 的属性集合 α_{k+1}^F 与 $\Delta \alpha_k^F$ 满足

$$(\alpha_k^F \cup \Delta \alpha_k^F) - \alpha_{k+1}^F = \emptyset \quad (38)$$

推论5 若 $w(x)_k^F$ 双依赖于 $w(x)_{k+1}^F, w(x)_{k+1}^F \Leftrightarrow w(x)_k^F$;则 $w(x)_k^F$ 的属性集合 α_k^F 与 $w(x)_{k+1}^F$ 的属性集合 α_{k+1}^F 满足

$$\alpha_k^F - \alpha_{k+1}^F = \emptyset \quad (39)$$

推论6 若 $w(x)_k^F, w(x)_t^F \in \{w(x)_i^F | i \in I\}, \tau < \lambda$,则

$$w(x)_k^F \Rightarrow w(x)_t^F \quad (40)$$

这里:式(40)中的 $w(x)_k^F, w(x)_t^F$ 分别是式(24)的表示。

推论7 若 $w(x)_k^F, w(x)_t^F \in \{w(x)_i^F | i \in I\}, w(x)_k^F \Rightarrow w(x)_t^F$;则 $w(x)_k^F$ 的属性集合 $\alpha_k^F, w(x)_t^F$ 的属性集合 α_t^F 满足

$$\alpha_k^F = \bigcup_{i \in I} \alpha_i^F \quad (41)$$

$$\alpha_t^F = \bigcap_{i \in I} \alpha_i^F \quad (42)$$

定理10(内P-信息规律单依赖链-属性单依赖链定理)

若 $w(x)_t^F, w(x)_j^F, w(x)_k^F \in \{w(x)_i^F | i \in T\}$ 构成单依赖链,而且

$$w(x)_k^F \Rightarrow w(x)_j^F \Rightarrow w(x)_t^F \quad (43)$$

则 $w(x)_k^F$ 的属性集合 $\alpha_k^F, w(x)_j^F$ 的属性集合 α_j^F 与 $w(x)_t^F$ 的属性集合 α_t^F 构成与 $w(x)_k^F \Rightarrow w(x)_j^F \Rightarrow w(x)_t^F$ 相反的属性单依赖链,而且

$$\alpha_k^F \Rightarrow \alpha_j^F \Rightarrow \alpha_t^F \quad (44)$$

推论8 若 $w(x)_t^F, w(x)_j^F, w(x)_k^F \in \{w(x)_i^F | i \in I\}$,而且

$$w(x)_k^F \Rightarrow w(x)_j^F \Rightarrow w(x)_t^F \quad (45)$$

则单依赖损失 $\nabla w(x)_j^F, \nabla w(x)_t^F$ 满足

$$\nabla w(x)_j^F \Rightarrow \nabla w(x)_t^F \quad (46)$$

这里:式(46)中, $\nabla w(x)_j^F$ 是 $w(x)_j^F$ 单依赖于 $w(x)_k^F$ 生成的单依赖损失, $\nabla w(x)_t^F$ 是 $w(x)_t^F$ 单依赖于 $w(x)_j^F$ 生成的单依赖损失。

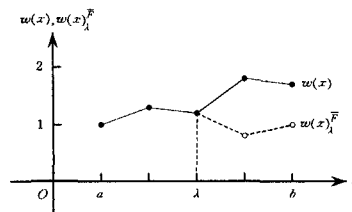
利用第3节、第4节中给出的讨论与结果,第5节中给出

如下。

5 内P-信息规律依赖生成的信息规律变轨与未知内P-信息规律依赖挖掘-发现

5.1 信息规律变轨直观表示与信息规律变轨特征

数学中,函数 $w(x)$ 是一条曲线(连续曲线或者离散曲线(折线)),信息科学与信息系统中,函数 $w(x)$ 被定义成信息规律 $w(x)^{[1,2]}$;图2给出信息规律 $w(x)$ 的变轨的直观表示。



$y = \{y_1, y_2, \dots, y_n\}$ 是函数集合S的离散数据集合, y 生成信息规律 $w(x)$ (折线), $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ 是 $w(x)$ 的属性集合。 $y^F = \{y_1^F, y_2^F, \dots, y_n^F\}$ 是函数内P-集合 S^F 的离散数据集合, y^F 生成内P-信息规律 $w(x)_\lambda^F$ (折线), $\alpha^F = \{\alpha_1, \alpha_2, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_r\}$ 是 $w(x)_\lambda^F$ 的属性集合。在 $[a, \lambda]$ 上, $w(x)$ 与 $w(x)_\lambda^F$ 重合;在 $[\lambda, b]$ 上, $w(x)_\lambda^F \leq w(x)$; $w(x)_\lambda^F$ 是 $w(x)$ 在 λ 点的变轨;在 λ 点, α 内被补充属性, α 生成 $\alpha^F, \alpha \subseteq \alpha^F$ 。

图2

由图2直接得到:

命题6 $w(x)_\lambda^F$ 是 $w(x)$ 在 $[\lambda, b]$ 上的变轨,必有 $\forall i \in [\lambda, b], y_i - y_i^F \geq 0$ 。这里: $y_i \in y, y_i^F \in y^F; i = \lambda, \lambda+1, \dots, b$ 。

命题7 $w(x)_\lambda^F$ 是 $w(x)$ 在 $[\lambda, b]$ 上的变轨,则 $w(x)_\lambda^F \Rightarrow w(x)$ 。

命题8 $\nabla w(x)$ 是 $[\lambda, b]$ 上 $w(x)$ 单依赖于 $w(x)_\lambda^F$ 生成的单依赖损失, $w(x)_\lambda^F$ 一定是 $w(x)$ 的变轨。

命题9 $w(x)_\lambda^F$ 是 $w(x)$ 在 $[\lambda, b]$ 上的变轨,一定存在单依赖损失 $\nabla w(x) \geq 0$ 。

5.2 未知内-P信息规律依赖挖掘-发现应用

本节例子取自风险投资利润规律估计,例子中的利润数据取自福建省龙岩市某国际投资集团公司 $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ 2010—2011年的利润年报, $\forall x_i \in X$ 是X的子公司, $i = 1, 2, 3, 4, 5, 6, 7; \alpha = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ 是X的市场属性(市场特征)。因为商业秘密原因,例子中X, α 的名称,略;例子中2010—2011年的利润数据已做了标准化,它们分别列入表1、表2;为了简单、直观,2010、2011年利润数据生成的利润规律用折线表示。

表1 集团公司X2010年1月—6月利润值分布, $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ 是X的市场属性集合

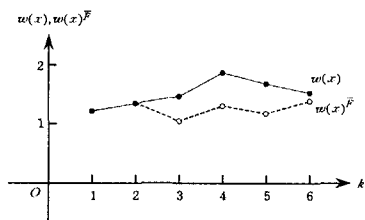
k	1	2	3	4	5	6
y	1.21	1.34	1.46	1.87	1.68	1.52

表2 集团公司X^F2011年1月—6月利润值分布, $\alpha^F = \alpha \cup \{\beta_1, \beta_2\}$ 是X^F的市场属性集合

k	1	2	3	4	5	6
y ^F	1.21	1.34	1.04	1.30	1.17	1.38

这里:表1中, $y = \{y_1, y_2, y_3, y_4, y_5, y_6\} = \{1.21, 1.34, 1.46, 1.87, 1.68, 1.52\}$ 是X在2010年1月—6月的利润值;表2中 $y^F = \{y_1^F, y_2^F, y_3^F, y_4^F, y_5^F, y_6^F\} = \{1.21, 1.34, 1.04, 1.30, 1.17, 1.38\}$ 是X^F在2011年1月—6月的利润值; $\forall x_i \in X$

(或者, $\forall x_i \in X^F$) 的利润值, 略。表 1 中 $y = \{\sum_{i=1}^7 y_{i,1}, \sum_{i=1}^7 y_{i,2}, \sum_{i=1}^7 y_{i,3}, \sum_{i=1}^7 y_{i,4}, \sum_{i=1}^7 y_{i,5}, \sum_{i=1}^7 y_{i,6}\}$ 。



集团公司 X2010 年 1 月—6 月的利润信息折线规律 $w(x)$, X 具有属性集合 α , $w(x)$ 用实线表示; 集团公司 X^F 2011 年 1 月—6 月的利润信息折线规律 $w(x)^F$, X^F 具有属性集合 α^F ; $w(x)^F$ 用虚线表示。 $w(x)$ 在 $k=2$ 出现变轨, $w(x)$ 变成 $w(x)^F$ 。

图 3

图 3 给出直接结论: 在离散区间 $[2, 6]$ 上, $w(x)$ 生成变轨, $w(x)$ 与 $w(x)^F$ 满足 $w(x)^F \Rightarrow w(x)$; 在 $\Delta\alpha = \{\beta_1, \beta_2\} \neq \emptyset$ 条件下, $w(x)^F$ 从 $w(x)$ 内被挖掘-发现; 在 $\Delta\alpha = \emptyset$ 的条件下, $w(x)^F$ 被潜藏在 $w(x)$ 内。

例子实证分析与认证

A. 2008 年年底爆发全球性金融危机, 金融危机对投资利润危害具有惰性与滞后性, 集团公司 X 在 2009 年的利润分布没有受到金融危机的冲击, 因为 2009 年的投资合同在 2008 年已被签约, 利润依约定支付。

B. 金融危机持续发作, 突发性的市场特征(市场属性) β_1 = “产品订单被取消”, β_2 = “结算货币贬值”随机生成, 稳定的市场属性集合 α 遭到 β_1, β_2 的入侵攻击, α 变成 α^F , $\alpha \subseteq \alpha^F = \alpha \cup \Delta\alpha = \alpha \cup \{\beta_1, \beta_2\}$; 在 α^F 存在的条件下, 子公司 x_2, x_6 相继被关闭(资不抵债), 集团公司 X 变成 $X^F = X - \{x_2, x_6\} = \{x_1, x_3, x_4, x_5, x_7\}$; 集团公司 X 在 2011 年 1 月—6 月的利润信息规律 $w(x)$ 在 $k=2$ 出现变轨, $w(x)$ 变成 $w(x)^F$, $w(x)^F \leq w(x)$ 。A 与 B 的结论与集团公司 X 的利润年报相吻合。在 $k=1, 2$; $w(x)$ 与 $w(x)^F$ 满足 $w(x) = w(x)^F$; 或者 $w(x)$ 与 $w(x)^F$ 重合是因为 β_1, β_2 入侵 α 内, 利润规律下跌具有滞后性, 不被立即表现。

C. 因为 $\Delta\alpha = \{\beta_1, \beta_2\} \neq \emptyset$; X 的属性集合 α , X^F 的属性集合 α^F 与 $\Delta\alpha$ 满足定理 3; 从图 3 中得到: $w(x)$ 与 $w(x)^F$ 满足定理 6, 而且 $w(x)^F \Rightarrow w(x)$; 从图 3 中得到: $k \geq 2$, 存在数据集 ∇y^F ; y, y^F 与 ∇y^F 满足定理 9; 在 $\alpha \subseteq \alpha^F$ 与 $k \geq 2$ 的条件下, $w(x)^F$ 从 $w(x)$ 内被挖掘-发现, 而且 $w(x)^F \Rightarrow w(x)$; 从图 3 中还得到: 在离散区间 $[2, 6]$ 上, $w(x)$ 与 $w(x)^F$ 满足命题 6—命题 9。折线利润信息规律 $w(x), w(x)^F$ 与 X 出具的利润年报相吻合。

结束语 P-集合是把动态特性引入到有限普通集合 X 内, 改进有限普通集合 X 提出的, P-集合被应用到一类信息系统中的动态信息特性与应用研究中^[4-22]; 因为我们遇到的信息具有动态特性的居多, 例如, 机器人对于运动目标的识别, 运动的目标就是一个动态信息。P-集合不能应用于运动目标的运动轨迹(规律)特征与应用研究中, 基于这个事实, 把函数概念引入到 P-集合内, 改进 P-集合, 提出函数 P-集合。区间(连续, 或者离散) $[a, b]$ 上的函数(连续, 或者离散)是 $[a, b]$ 上的一个规律(连续, 或者离散)。本文利用函数 P-集合的结构与依赖概念交叉给出内 P-信息规律依赖内-挖掘、信息规律变轨概念、内 P-信息规律依赖内-挖掘的应用。函数 P-集合是研究动态信息规律(轨迹)特性与应用的一个新方法、新

模型。信息规律变轨是函数 P-集合的一个重要应用特性。

参考文献

- [1] 史开泉. 函数 P-集合[J]. 山东大学学报: 理学版, 2011, 46(2): 62-69
- [2] Shi Kai-quan. Function P-sets [J]. International Journal of Machine Learning and Cybernetics, 2011, 2(4): 281-288
- [3] 赵文菊, 范成贤. 函数 P-集合属性依赖与应用[J]. 山东大学学报: 理学版, 2011, 46(6): 115-126
- [4] 史开泉. P-集合[J]. 山东大学学报: 理学版, 2008, 43(11): 78-84
- [5] Shi Kai-quan. P-sets and its applications [J]. An International Journal Advances in Systems Science and Applications, 2009, 9(2): 209-219
- [6] 史开泉. P-集合与它的应用特性[J]. 计算机科学, 2010, 37(8): 1-8
- [7] 史开泉. P-推理与信息的 P-推理发现-辨识[J]. 计算机科学, 2011, 38(7): 1-9
- [8] 史开泉. P-集合, 逆 P-集合与信息智能融合-过滤辨识[J]. 计算机科学, 2012, 39(4): 1-13
- [9] 李豫颖, 范成贤, 史开泉. 混合记忆信息与记忆筛选[J]. 系统工程与电子技术, 2011, 33(8): 1824-1828
- [10] 赵树理, 范成贤, 史开泉. 倒向 P-推理与内 P-属性类特征[J]. 计算机科学, 2012, 39(7): 245-249
- [11] 赵树理, 范成贤, 史开泉. 外 P-信息生成与它的推理-搜索发现[J]. 山东大学学报: 理学版, 2012, 47(11): 99-104
- [12] Fan Cheng-xian, Lin Hong-kang. P-sets and the reasoning-identification of disaster information [J]. An International Journal of Convergence Information Technology, 2012, 7(1): 337-345
- [13] Lin Hong-kang, Fan Cheng-xian. The dual form of P-reasoning and identification of unknown attribute[J]. International Journal of Digital Content Technology and its Applications, 2012, 6(1): 121-131
- [14] Lin Rong, Fan Cheng-xian. P-sets and identification of inward-convergence information[J]. An International Journal of Convergence Information Technology, 2012, 7(7): 157-164
- [15] Lin Hong-kang, Fan Cheng-xian. Embedding-camouflage of inverse P-information and applications[J]. An International Journal of Convergence Information Technology, 2012, 7(20): 471-480
- [16] Fan Cheng-xia, Huang Shun-liang. Inverse P-reasoning discovery identification of inverse P- information[J]. International Journal of Digital Content Technology and its Applications, 2012, 6(20): 735-744
- [17] Zhang Li, Cui Yu-quan. Outer P-sets and data internal recovery [J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 189-199
- [18] 阎红灿, 王坚, 刘保相. 基于 P-集合的本体形式背景抽取[J]. 计算机应用研究, 2012, 29(6): 2196-2204
- [19] 耿红琴, 张冠宇, 史开泉. F-信息伪装与伪装-还原辨识[J]. 计算机科学, 2011, 38(2): 241-245
- [20] 李豫颖, 林宏康, 史开泉. 数据离散区间特征与数据发现-应用[J]. 系统工程与电子技术, 2011, 33(10): 2258-2262
- [21] 于秀清. 迭代 F-内嵌入信息生成及其遗传发现-应用[J]. 系统工程与电子技术, 2011, 33(12): 2691-2731
- [22] 周玉华, 张冠宇, 史开泉. P-集合与双信息规律生成[J]. 数学的实践与认识, 2010, 40(13): 71-80