基于单边区间集概念格的不完备形式背景的属性约简

王 振 魏 玲
（西北大学数学学院 西安 710127）

摘 要 单边区间集概念的提出为不完备形式背景的数据分析奠定了理论基础，也为研究其属性约简提供了思路。首先给出了不完备形式背景上的 4 种约简，即保持单边区间集概念格结构不变的约简、保持并(交)不可约元外延不变的约简与保持对象单边区间集概念外延不变的约简，并研究了它们的关系，最后给出了基于差别矩阵与差别函数计算约简的方法。

关键词 不完备形式背景，单边区间集概念，属性约简，差别矩阵
中图法分类号 O29，TP181 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.01.011

Attribute Reduction of Partially-known Formal Concept Lattices for Incomplete Contexts

WANG Zhen WEI Ling
(Department of Mathematics, Northwest University, Xi'an 710127, China)

Abstract Partially-known formal concept, which was proposed recently, lays the foundation of data analysis of incomplete contexts and also provides the thought of studying on attribute reduction. This paper firstly proposed four kinds of attribute reduction: partially-known formal concept lattice reduction, meet(join)-irreducible elements preserving reduction and partially-known object formal concept preserving reduction. And then, it discussed the relationships among the four kinds of reduction. Finally, it presented the approaches to finding these reduction by discernibility matrices and discernibility functions.

Keywords Incomplete context, Partially-known formal concept, Attribute reduction, Discernibility matrix

1 引言

形式概念分析[1]是 Wille 教授于 1982 年提出的一种从形式背景中进行数据分析和规则提取的强有力工具，其理论基础为形式概念。概念格是形式概念分析中的核心数据结构，本质上描述了对象与属性之间的联系。

作为数据分析及知识处理的有力工具，概念格理论已经被广泛应用于医学[2]、概念格的数据挖掘[3-4]、信息检索及软件工程[5-7]等领域。随着形式概念分析的发展，国内外许多学者对其进行了深入的研究，并在概念格的构造与属性约简[8-11]、规则提取[12]、向三元概念分析的拓广[13-14]等方面取得了一定的成果。

形式背景是形式概念分析的基础研究对象，由于信息缺失的发生，Burmeister 等人在形式背景的基础上提出了不完备形式背景[15]。针对不完备形式背景，李金海等人定义了近似概念并构建了近似概念格[16]；Yao 根据区间集理论，提出了不完备形式背景上的三支概念分析[17]；李美争等人[18]根据三支概念分析构建了三支近似概念并研究了基于三支近似概念格的不完备形式背景属性约简。由于不完备形式背景对实际的贴合性，因此基于不完备形式背景的形式概念分析将成为数据分析和规则提取的实用工具。

单边区间集概念格是不完备形式背景的概念分析的基础，从不完备形式背景获取的知识间的关系都会反映在单边区间集概念格上，基于此，本文提出保持单边区间集概念格结构不变的约简；另一方面，根据单边区间集概念中的交(并)不可约元是下(上)确界稠密的，提出了保持不可约元外延不变的约简来保证格构建中的基本元。从粒计算的角度出发，在单边区间集概念格中每一个对象的单边区间集概念都可被视为信息粒度，因此提出保持对象单边区间集概念外延不变的约简，并进一步研究了它们的关系，给出了各种约简的计算方法。

2 预备知识

本节给出所需的基于不完备形式背景的相关概念。

定义 1[15] 称四元对$(U,A,\{+,?,-\},I)$为一个不完备形式背景，其中 U 为对象集，A 为属性集，$\{+,?,-\}$为集值，$I\subseteq U\times A\times\{+,?,-\}$为 $U,A,\{+,?,-\}$间的三元关系，使得$(x,a,+)$表示对象 x 拥有属性 a，$(x,a,-)$表示对象 x 不拥有属性 a，$(x,a,?)$表示不确定对象 x 是否拥有属性 a。通常记$\{+,?,-\}$为 V，故不完备形式背景$(U,A,\{+,?,-\},I)$可记为(U,A,V,I)。

到稿日期：2017-03-03 返修日期：2017-05-14 本文受国家自然科学基金项目(11371014，11071281)资助。
王 振(1992—)，男，硕士生，主要研究方向为粗糙集、形式概念分析等；魏 玲(1972—)，女，博士生，教授，CCF 会员，主要研究方向为粗糙集、形式概念分析、概率论等，E-mail：wl@nwu.edu.cn(通信作者)。

例 1　不完备形式背景(U,A,V,I)如表 1 所列，集合$U=\{1,2,3,4\}$表示 4 位病患，集合 $A=\{a,b,c,d,e,f,g\}$表示 7 种症状，若第一位病患有症状 a，则表中对应位置用＋表示；第二位病患不确定是否患有症状 a，则表中对应位置用？表示；否则用－表示。

表 1　不完备形式背景(U,A,V,I)

Table 1　Incomplete formal context(U,A,V,I)

	发烧	咳嗽	头疼	呼吸困难	腹泻	耳鸣	易困
1	＋	＋	－	＋	＋	－	＋
2	?	?	＋	－	－	?	?
3	－	－	－	＋	－	－	?
4	＋	＋	＋	－	－	＋	－

设 A 为属性集，$2^A\times2^A$ 为 2^A 和 2^A 间的笛卡尔积，$\forall[B_1,C_1],[B_2,C_2]\in2^A\times2^A$，定义其偏序关系$\leqslant$为：$[B_1,C_1]\leqslant[B_2,C_2]\Leftrightarrow B_1\subseteq B_2$ 且 $C_1\subseteq C_2$。特别地，若 $B_1=B_2$ 且 $C_1=C_2$，则$[B_1,C_1]=[B_2,C_2]$，其交($\cap$)、并($\cup$)、差($-$)运算为：$[B_1,C_1]\cap[B_2,C_2]=[B_1\cap B_2,C_1\cap C_2]$，$[B_1,C_1]\cup[B_2,C_2]=[B_1\cup B_2,C_1\cup C_2]$，$[B_1,C_1]-[B_2,C_2]=[B_1-B_2,C_1-C_2]$。

对于不完备形式背景(U,A,V,I)，$X\subseteq U$，定义：

$\underline{R}(X)=\{a\in A|\forall x\in X,I(x,a)=+\}$

$\bar{R}(X)=\{a\in A|\forall x\in X,I(x,a)=+$或$I(x,a)=?\}$

$\underline{R}(X)$表示 X 中对象共同确定拥有的属性集，$\bar{R}(X)$表示 X 中对象共同可能拥有的属性集，显然$\underline{R}(X)\subseteq\bar{R}(X)$。

定义 2[16]　设(U,A,V,I)为不完备形式背景，对 $X\in2^U$，$[B,C]\in2^A\times2^A$ 分别定义运算：$X^{\lhd}=[\underline{R}(X),\bar{R}(X)]$，$[B,C]^{\rhd}=\{x\in U|[B,C]\leqslant[\underline{R}(x),\bar{R}(x)]$。

定义 3[16]　设(U,A,V,I)为不完备形式背景，$S\subseteq A$，称(U,S,V,I_S)为(U,A,V,I)的子背景，其中 $I_S=I\cap U\times S$。

子背景上的$\lhd$，$\rhd$运算定义为：$X^{\lhd_S}=[\underline{R}_S(X),\bar{R}_S(X)]$，$[B,C]^{\rhd_S}=\{x\in U|[B,C]\leqslant[\underline{R}_S(x),\bar{R}_S(x)]\}$。其中，$\underline{R}_S(X)=\{a\in S\mid\forall x\in X,I_S(x,a)=+\}$，$\bar{R}_S(X)=\{a\in S|\forall x\in X,I_S(x,a)=+$或$I(x,a)=?\}$。易知，$\lhd_A=\lhd$，$\rhd_A=\rhd$。

由定义 2 与定义 3 可得算子性质如下。

命题 1[16]　设(U,A,V,I)为不完备形式背景，对于 $S\subseteq A,X,X_1,X_2\in2^U$，$[B,C]$，$[B_1,C_1]$，$[B_2,C_2]\in2^A\times2^A$，下列性质成立：

(1)$X_1\subseteq X_2\Rightarrow X_2^{\lhd}\leqslant X_1^{\lhd}$，$[B_1,C_1]\leqslant[B_2,C_2]\Rightarrow[B_2,C_2]^{\rhd}\subseteq[B_1,C_1]^{\rhd}$；

(2)$X\subseteq X^{\lhd\rhd}$，$[B,C]\leqslant[B,C]^{\rhd\lhd}$；

(3)$X^{\lhd}=X^{\lhd\rhd\lhd}$，$[B,C]^{\rhd}=[B,C]^{\rhd\lhd\rhd}$；

(4)$(X_1\cup X_2)^{\lhd}=X_1^{\lhd}\cap X_2^{\lhd}$，$([B_1,C_1]\cap[B_2,C_2])^{\rhd}=[B_1,C_1]^{\rhd}\cup[B_2,C_2]^{\rhd}$；

(5) $X^{\lhd_S}=X^{\lhd_A}\cap[S,S]$，$[B,C]^{\rhd_S}=[B,C]^{\rhd}$。

定义 4[16]　设(U,A,V,I)为不完备形式背景，对于 $X\in2^U$，$[B,C]\in2^A\times2^A$，若 $X^{\lhd}=[B,C]$，$[B,C]^{\rhd}=X$，则称$(X,[B,C])$为(U,A,V,I)的单边区间集概念。

本文研究了当 X 为单点集$\{x\}$时的单边区间集概念$(x^{\lhd\rhd},x^{\lhd})$，我们称之为对象单边区间集概念。

命题 2[16]　设(U,A,V,I)为不完备形式背景，则单边区间集概念集$\{(X_t,[B_t,C_t])|t\in T\}$的上($\vee$)、下($\wedge$)确界如下：

$$\bigwedge_{t\in T}(X_t,[B_t,C_t])=(\bigcap_{t\in T}X_t,(\bigcup_{t\in T}[B_t,C_t])^{\rhd\lhd})$$

$$\bigvee_{t\in T}(X_t,[B_t,C_t])=((\bigcup_{t\in T}X_t)^{\lhd\rhd},\bigcap_{t\in T}[B_t,C_t])$$

用 $S\text{-}ISL(U,A,V,I)$表示所有单边区间集概念的集合，记$(X_1,[B_1,C_1])\leqslant(X_2,[B_2,C_2])\Leftrightarrow X_1\subseteq X_2\Leftrightarrow[B_2,C_2]\leqslant[B_1,C_1]$，则“$\leqslant$”是 $S\text{-}ISL(U,A,V,I)$上的偏序关系，称$(X_2,[B_2,C_2])$为$(X_1,[B_1,C_1])$的超单边区间集概念，$(X_1,[B_1,C_1])$为$(X_2,[B_2,C_2])$的亚单边区间集概念。进一步，若不存在$(X_3,[B_3,C_3])$，使得$(X_1,[B_1,C_1])<(X_3,[B_3,C_3])<(X_2,[B_2,C_2])$，其中$(X_1,[B_1,C_1])<(X_3,[B_3,C_3])\Leftrightarrow(X_1,[B_1,C_1])\leqslant(X_3,[B_3,C_3])$且$(X_1,[B_1,C_1])\neq(X_3,[B_3,C_3])$，则称$(X_2,[B_2,C_2])$为$(X_1,[B_1,C_1])$的父单边区间集概念，$(X_1,[B_1,C_1])$为$(X_2,[B_2,C_2])$的子单边区间集概念，记为$(X_1,[B_1,C_2])\prec(X_2,[B_2,C_2])$。由命题 2 可知，$S\text{-}ISL(U,A,V,I)$是完备格，称为不完备形式背景$(U,A,V,I)$的单边区间集概念格。

例 2(续例 1)　不完备形式背景(U,A,V,I)的单边区间集概念格如图 1 所示。

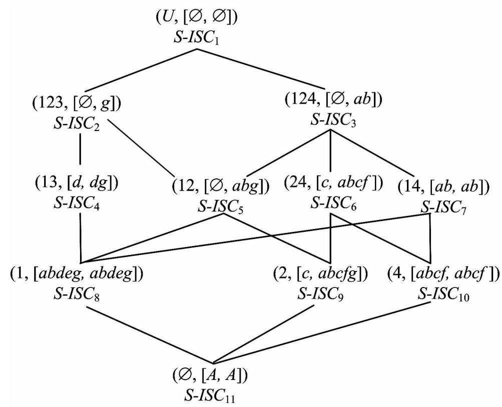

图 1　$S\text{-}ISL(U,A,V,I)$

Fig. 1　$S\text{-}ISL(U,A,V,I)$

定义 5[19]　设 L 是一个格，$\forall x\in L$，称 x 为一个并不可约元，若：1)$x\neq0$(如果 L 有零元)；2)$\forall a,b\in L,x=a\vee b\Rightarrow x=a$或$x=b$。对偶地，交不可约元可以同样地定义。

3　单边区间集概念格上的 4 种约简

本节从单边区间集概念格结构、单边区间集概念格构建、粒计算等角度出发，给出了不完备形式背景上的 4 种约简。

定义 6　设(U,A,V,I)为不完备形式背景，$D\subseteq A$，若 $Ext_{S\text{-}ISL}(U,A,V,I)=Ext_{S\text{-}ISL}(U,D,V,I_D)$，其中 $I_D=I\cap U\times D$，$Ext_{S\text{-}ISL}(U,A,V,I)=\{X|(X,[B,C])\in S\text{-}ISL(U,A,V,I)\}$，则称 D 为保持单边区间集概念格结构不变的协调集，简称为格协调集。进一步，若 $Ext_{S\text{-}ISL}(U,A,V,I)\neq Ext_{S\text{-}ISL}(U,D-\{d\},V,I_{D-\{d\}})$ $(\forall d\in D)$，则称 D 为保持单边区间集概念格结构不变的约简，简称为格约简。

在有限格中，交不可约元集在格中是下确界稠密的，并不可约元集是上确界稠密的，因此交(并)不可约元为格构建的基本元，基于此定义了保持交(并)不可约元外延不变的约简。

定义 7　设 (U,A,V,I) 为不完备形式背景,$D\subseteq A$,若 $Ext_{S\text{-}ISM}(U,A,V,I)=Ext_{S\text{-}ISM}(U,D,V,I_D)$,其中 $I_D=I\cap U\times D$,$Ext_{S\text{-}ISM}(U,A,V,I)=\{X\mid(X,[B,C])\in S\text{-}ISL(U,A,V,I)$ 且 $(X,[B,C])$ 为交不可约元$\}$,则称 D 为保持交不可约元外延不变的协调集。进一步,若 $Ext_{S\text{-}ISM}(U,A,V,I)\neq Ext_{S\text{-}ISM}(U,D-\{d\},V,I_{D-\{d\}})(\forall d\in D)$,则称 D 为保持交不可约元外延不变的约简。

定义 8　设 (U,A,V,I) 为不完备形式背景,$D\subseteq A$,若 $Ext_{S\text{-}ISJ}(U,A,V,I)=Ext_{S\text{-}ISJ}(U,D,V,I_D)$,其中 $I_D=I\cap U\times D$,$Ext_{S\text{-}ISJ}(U,A,V,I)=\{X\mid(X,[B,C])\in S\text{-}ISL(U,A,V,I)$ 且 $(X,[B,C])$ 为并不可约元$\}$,则称 D 为保持并不可约元外延不变的协调集。进一步,若 $Ext_{S\text{-}ISJ}(U,A,V,I)\neq Ext_{S\text{-}ISJ}(U,D-\{d\},V,I_{D-\{d\}})(\forall d\in D)$,则称 D 为保持并不可约元外延不变的约简。

与经典形式概念分析类似,每一个单边区间集概念都可被对象单边区间集概念的并给出,因此每一个对象单边区间集概念可被视为信息粒度,从粒计算的角度出发,提出保持对象单边区间集概念外延不变的约简。

定义 9　设 (U,A,V,I) 为不完备形式背景,$D\subseteq A$,若 $\forall x\in U,x^{\lhd_A\rhd_A}=x^{\lhd_D\rhd_D}$,则称 D 为保持对象单边区间集概念外延不变的协调集。进一步,若 $\exists x\in U$,使得 $x^{\lhd\rhd}\neq x^{\lhd_{D-\{d\}}\rhd_{D-\{d\}}}(\forall d\in D)$,则称 D 为保持对象单边区间集概念外延不变的约简。

按照提出的顺序,把 4 种协调集的集合依次记为 $CS(S\text{-}ISL)$,$CS(S\text{-}ISM)$,$CS(S\text{-}ISJ)$,$CS(S\text{-}ISG)$。4 种约简的集合依次记为 $RED(S\text{-}ISL)$,$RED(S\text{-}ISM)$,$RED(S\text{-}ISJ)$,$RED(S\text{-}ISG)$。

下面给出 4 种约简间的关系,首先给出引理 1。

引理 1[19]　设 L 为一个有限格,则 L 中的任意一个元素可以表示为 L 中并(交)不可约元的并(交)。

定理 1　设 (U,A,V,I) 为不完备形式背景,则 $CS(S\text{-}ISL)=CS(S\text{-}ISM)$,$RED(S\text{-}ISL)=RED(S\text{-}ISM)$。

证明:首先证明 $CS(S\text{-}ISL)=CS(S\text{-}ISM)$。若 $D\in CS(S\text{-}ISL)$,则 $Ext_{S\text{-}ISL}(U,A,V,I)=Ext_{S\text{-}ISL}(U,D,V,I_D)$ 且 $Ext_{S\text{-}ISL}(U,A,V,I)$,$Ext_{S\text{-}ISL}(U,D,V,I_D)$ 在集合包含关系下形成格。$\forall X\in Ext_{S\text{-}ISM}(U,A,V,I)$,由交不可约元的定义可得,$\forall Y,Z\in Ext_{S\text{-}ISL}(U,A,V,I)$,$X\neq Y$,$X\neq Z$,则 $X\neq Y\cap Z$。由 $Ext_{S\text{-}ISL}(U,A,V,I)=Ext_{S\text{-}ISL}(U,D,V,I_D)$,则 $X,Y,Z\in Ext_{S\text{-}ISL}(U,D,V,I_D)$,且 $X\neq Y$,$X\neq Z$,则 $X\neq Y\cap Z$。那么由交不可约元的定义可得,$X\in Ext_{S\text{-}ISM}(U,D,V,I_D)$,因此 $Ext_{S\text{-}ISM}(U,A,V,I)\subseteq Ext_{S\text{-}ISM}(U,D,V,I_D)$。反之亦成立,因此 $Ext_{S\text{-}ISM}(U,A,V,I)\subseteq Ext_{S\text{-}ISM}(U,D,V,I_D)$,即 $D\in CS(S\text{-}ISM)$。

若要 $D\in CS(S\text{-}ISM)$,需证 $Ext_{S\text{-}ISL}(U,A,V,I)=Ext_{S\text{-}ISL}(U,D,V,I_D)$。由 $Ext_{S\text{-}ISL}(U,D,V,I_D)\subseteq Ext_{S\text{-}ISL}(U,A,V,I)$,只需证 $Ext_{S\text{-}ISL}(U,A,V,I)\subseteq Ext_{S\text{-}ISL}(U,D,V,I_D)$。由引理 1,$\forall X\in Ext_{S\text{-}ISL}(U,A,V,I)$,$\exists X_i\in Ext_{S\text{-}ISM}(U,A,V,I)$,使得 $X=\bigcap_{i\in I}X_i$。由 $D\in CS(S\text{-}ISM)$,则 $Ext_{S\text{-}ISM(U,A,V,I)=}Ext_{S\text{-}ISM}(U,D,V,I_D)$。又 $Ext_{S\text{-}ISM}(U,D,V,I_D)\subseteq Ext_{S\text{-}ISL}(U,D,V,I_D)$,因此 $X=\bigcap_{i\in I}X_i\in Ext_{S\text{-}ISL}(U,D,V,I_D)$。故 $Ext_{S\text{-}ISL}(U,A,V,I)\subseteq Ext_{S\text{-}ISL}(U,D,V,I_D)$,则 $Ext_{S\text{-}ISL}(U,A,V,I)=Ext_{S\text{-}ISL}(U,D,V,I_D)$,即 $D\in CS(OEL)$。综上所述,$CS(S\text{-}ISL)=CS(S\text{-}ISM)$。同理 $RED(S\text{-}ISL)=RED(S\text{-}ISM)$。

定理 1 表明在不完备形式背景中,保持单边区间集概念格结构不变的约简(协调集)与保持交不可约元外延不变的约简(协调集)是等价的。

引理 2　设 (U,A,V,I) 为不完备形式背景,$X\subseteq U$ 及 $D\subseteq A$。若 D 为一个协调集,则 $X^{\lhd_A\rhd_A}=X^{\lhd_D\rhd_D}$。

证明:$\forall X\subseteq U$,由 $D\subseteq A$ 和子背景算子性质,$X^{\lhd_D}\subseteq X^{\rhd_A}$。假设 D 是一个协调集,$\forall X\subseteq U$,由 $X^{\lhd_D\rhd_D}\in Ext_{S\text{-}ISL}(U,D,V,I_D)$ 及 $X^{\lhd_A\rhd_A}\in Ext_{S\text{-}ISL}(U,A,V,I)$,则 $X^{\lhd_A\rhd_A}=X^{\lhd_D\rhd_D}$。

定理 2　设 (U,A,V,I) 为不完备形式背景,则 $CS(S\text{-}ISL)\subseteq CS(S\text{-}ISG)$,$RED(S\text{-}ISL)\subseteq CS(S\text{-}ISG)$。

证明:由引理 2,$CS(S\text{-}ISL)\subseteq CS(S\text{-}ISG)$。又 $RED(S\text{-}ISL)\subseteq CS(S\text{-}ISL)$,则 $RED(S\text{-}ISL)\subseteq CS(S\text{-}ISG)$。

根据定理 2,保持单边区间集概念格结构不变的协调集(约简)一定为保持对象单边区间集概念外延不变的协调集。

引理 3　设 (U,A,V,I) 为不完备形式背景,$(X,[B,C])$ 为单边区间集概念,若 $(X,[B,C])$ 为 $S\text{-}ISL(U,A,V,I)$ 中的并不可约元,则 $(X,[B,C])$ 为对象单边区间集概念。

证明:设 $(X,[B,C])\in S\text{-}ISL(U,A,V,I)$,则 $\bigvee_{x_i\in X}(x_i^{\lhd\rhd},x_i^{\lhd})=\bigvee_{x_i\in X}(x_i^{\lhd\rhd},[\underline{R}(x_i),\bar{R}(x_i)])=((\bigcup_{x_i\in X}x_i)^{\lhd\rhd},[\bigcap_{x_i\in X}\underline{R}(x_i),\bigcap_{x_i\in X}\bar{R}(x_i)])$。由 $\underline{R}$ 和 $\bar{R}$ 的性质,$\bigcap_{x_i\in X}\underline{R}(x_i)=\underline{R}(\bigcup_{x_i\in X}x_i)=\underline{R}(X)=A$,$\bigcap_{x_i\in X}=\bar{R}(x_i)=\underline{R}(\bigcup_{x_i\in X}x_i)=\bar{R}(X)=B$,因此 $(X,[B,C])=\bigvee_{x_i\in X}(x_i^{\lhd\rhd},x_i^{\lhd})$。故 $\exists x_j\in X$,使得 $(X,[B,C])=(x_j^{\lhd\rhd},x_j^{\lhd})$,因此并不可约元为对象单边区间集概念。

定理 3　设 (U,A,V,I) 为不完备形式背景,则 $CS(S\text{-}ISL)\subseteq CS(S\text{-}ISJ)$,$RED(S\text{-}ISL)\subseteq CS(S\text{-}ISJ)$。

证明:由引理 3,显然成立。

由定理 3 可知,保持单边区间集概念格结构不变的约简(协调集)一定为保持并不可约元外延不变的协调集。

由上述分析得到 4 种约简间的关系,如图 2 所示。

D 为保持并不可约元外延不变的协调集

$\Uparrow$

D 为保持单边区间集概念格结构不变的协调集

$\Updownarrow$

D 为保持交不可约元外延不变的协调集

$\Downarrow$

D 为保持对象单边区间集概念外延不变的协调集

图 2　4 种协调集的关系

Fig. 2　Relationships among four kinds of consistent sets

4　属性约简的方法

本节给出不完备形式背景的属性约简的计算方法。为此,首先基于张文修等人在经典形式概念分析中提出的差别矩阵[20]给出了不完备形式背景下的差别矩阵。

定义 10　设 (U,A,V,I) 为不完备形式背景,$(X,[B,

$C]$)，$(Y,[E,F])\in S\text{-}ISL(U,A,V,I)$，则

$$DIS_{S\text{-}ISL}((X,[B,C]),(Y,[E,F]))=\begin{cases}(B-E)\cup(C-F), & (X,[B,C])\prec(Y,[E,F])\\ \emptyset, & \text{其他}\end{cases}$$

称为$(X,[B,C])$，$(Y,[E,F])$间的差别属性集。

记 $\bigwedge_{S\text{-}ISL}=(DIS_{S\text{-}ISL}((X,[B,C]),(Y,[E,F])))$为$(U,A,V,I)$的 S-ISL-差别矩阵，其中$(X,[B,C])$，$(Y,[E,F])\in S\text{-}ISL(U,A,V,I)$。特别地，若$(X,[B,C])$为交不可约元，则称之为$(U,A,V,I)$的 S-ISM-差别矩阵，记为 $\bigwedge_{S\text{-}ISM}$；若$(X,[B,C])$为并不可约元，则称之为(U,A,V,I)的 S-ISJ-差别矩阵，记为 $\bigwedge_{S\text{-}ISJ}$；若$(X,[B,C])$为对象单边区间集概念，则称之为(U,A,V,I)的 S-ISG-差别矩阵，记为 $\bigwedge_{S\text{-}ISG}$。

由于在属性约简的计算过程中仅需要非空集，因此用$\wedge$来记非空差别属性集的集合。

引理 4 设(U,A,V,I)为不完备形式背景，$D\subseteq A$，$(X_i,[B_i,C_i])$，$(X_j,[B_j,C_j])$，$(X_k,[B_k,C_k])$为单边区间集概念，且$(X_i,[B_i,C_i])\prec(X_j,[B_j,C_j])\leqslant(X_k,[B_k,C_k])$，若$[B_i,C_i]\cap[D,D]\neq[B_j,C_j]\cap[D,D]$，则$[B_i,C_i]\cap[D,D]\neq[B_k,C_k]\cap[D,D]$。

证明：由$(X_i,[B_i,C_i])\prec(X_j,[B_j,C_j])$，则$[B_j,C_j]<[B_i,C_i]$，则$[B_j,C_j]\cap[D,D]\leqslant[B_i,C_i]\cap[D,D]$。又$[B_i,C_i]\cap[D,D]\neq[B_j,C_j]\cap[D,D]$，故$[B_j,C_j]\cap[D,D]<[B_i,C_i]\cap[D,D]$。因为$(X_j,[B_j,C_j])\leqslant(X_k,[B_k,C_k])$，所以$[B_k,C_k]\leqslant[B_j,C_j]$，故$[B_k,C_k]\cap[D,D]\leqslant[B_j,C_j]\cap[D,D]$。又$[B_j,C_j]\cap[D,D]<[B_i,C_i]\cap[D,D]$，因此$[B_i,C_i]\cap[D,D]\neq(B_k,C_k)\cap[D,D]$。

定理 4 设(U,A,V,I)为不完备形式背景，$D\subseteq A$，$D\neq\emptyset$，则下列论述等价：

(1)D为格协调集；

(2)$\forall(X_i,[B_i,C_i])$，$(X_j,[B_j,C_j])\in S\text{-}ISL(U,A,V,I)$，若$(X_i,[B_i,C_i])\prec(X_j,[B_j,C_j])$，则$[B_i,C_i]\cap[D,D]\neq[B_j,C_j]\cap[D,D]$；

(3)$\forall(X,[B,C])\in S\text{-}ISL(U,A,V,I)$，$(Y,[E,F])\in PC((X,[B,C]))$，则$D\cap DIS_{S\text{-}ISL}((X,[B,C]),(Y,[E,F]))\neq\emptyset$，其中$PC((X,[B,C]))$为$(X,[B,C])$的父概念集。

证明：先证(1)$\Leftrightarrow$(2)。

(1)$\Rightarrow$(2)。设$(X_i,[B_i,C_i])$，$(X_j,[B_j,C_j])\in S\text{-}ISL(U,A,V,I)$且$(X_i,[B_i,C_i])\prec(X_j,[B_j,C_j])$。因为$D$为格协调集，所以$S\text{-}ISL(U,D,V,I_D)\leqslant S\text{-}ISL(U,A,V,I)$，故存在$E_i,E_j,F_i,F_j$，使得$(X_i,[E_i,F_i])$，$(X_j,[E_j,F_j])\in S\text{-}ISL(U,D,V,I_D)$。由假设$(X_i,[B_i,C_i])\prec(X_j,[B_j,C_j])$可得$X_i\neq X_j$，因此$[E_i,F_i]\neq[E_j,F_j]$。根据$\lhd$的性质，$[E_i,F_i]=X_i^{\lhd_D}=X_i^{\lhd_A}\cap[D,D]=[B_i,C_i]\cap[D,D]$，同理可得$[E_j,F_j]=X_j^{\lhd_D}=X_j^{\lhd_A}\cap[D,D]=[B_j,C_j]\cap[D,D]$，因此$[B_i,C_i]\cap[D,D]\neq[B_j,C_j]\cap[D,D]$。

(2)$\Rightarrow$(1)。设$D\subseteq A$，$D\neq\emptyset$，则$S\text{-}ISL(U,A,V,I)\leqslant S\text{-}ISL(U,D,V,I_D)$。故要证$D$为格协调集，只需证$S\text{-}ISL(U,D,V,I_D)\leqslant S\text{-}ISL(U,A,V,I)$，即$\forall(X,[B,C])\in S\text{-}ISL(U,A,V,I)$，$(X,([B,C]\cap[D,D]))\in S\text{-}ISL(U,D,V,I_D)$。因此需证$X^{\lhd_D}=[B,C]\cap[D,D]$，$([B,C]\cap[D,D])^{\rhd_D}=X$。又因为$X^{\lhd_D}=X^{\lhd_A}\cap[D,D]=[B,C]\cap[D,D]$，所以只需证$([B,C]\cap[D,D])^{\rhd_D}=X$。假设$([B,C]\cap[D,D])^{\rhd_D}\neq X$，由$(([B,C]\cap[D,D])^{\rhd_A},([B,C]\cap[D,D])^{\rhd_A\lhd_A})\in S\text{-}ISL(U,A,V,I)$，则$(X,[B,C])\prec((([B,C]\cap[D,D])^{\rhd_A},([B,C]\cap[D,D])^{\rhd_A\lhd_A})$。因此存在$(Y,[E,F])\in S\text{-}ISL(U,A,V,I)$，使得$(X,[B,C])\prec(Y,[E,F])\leqslant(([B,C]\cap[D,D])^{\rhd_A},([B,C]\cap[D,D])^{\rhd_A\lhd_A})$，故$[B,C]\cap[D,D]\neq[E,F]\cap[D,D]$。由引理 4，可得$[B,C]\cap[D,D]\neq([B,C]\cap[D,D])^{\rhd_A\lhd_A}\cap[D,D]$。又因为$([B,C]\cap[D,D])^{\rhd_A\lhd_A}<[B,C]$且$[B,C]\cap[D,D]\leqslant([B,C]\cap[D,D])^{\rhd_A\lhd_A}\Rightarrow[B,C]\cap[D,D]=[B,C]\cap[D,D]\cap[D,D]\leqslant([B,C]\cap[D,D])^{\rhd_A\lhd_A}\cap[D,D]$，故$[B,C]\cap[D,D]=([B,C]\cap[D,D])^{\rhd_A\lhd_A}\cap[D,D]$，矛盾，所以$([B,C]\cap[D,D])^{\rhd_D}=X$。

下面证明：(1)$\Leftrightarrow$(3)。

(1)$\Rightarrow$(3)。假设$(X,[B,C])=(X,X^{\lhd_A})\in S\text{-}ISL(U,A,V,I)$，$(Y,[E,F])\in PC((X,[B,C]))$。由于$D$为格协调集，因此$X^{\lhd_D\rhd_D}=X$。

①若$(Y,([E,F]\cap[D,D]))\in S\text{-}ISL(U,A,V,I)$，则$[E,F]\cap[D,D]=[E,F]$。又$(X,[B,C])\prec(Y,[E,F])$可得$[E,F]\cap[D,D]<[B,C]\cap[D,D]$，则$[B,C]\cap[D,D]$-$[E,F]\cap[D,D]=[(B\text{-}E)\cap D,(C\text{-}F)\cap D]\neq[\emptyset,\emptyset]$，故$D\cap DIS_{S\text{-}ISL}((X,[B,C]),(Y,[E,F]))\neq\emptyset$。

②若$(Y,[E,F]\cap[D,D])\notin S\text{-}ISL(U,A,V,I)$，则$X\subset Y\subset([E,F]\cap[D,D])^{\rhd_D}$。又因为$(([E,F]\cap[D,D])^{\rhd_D},([E,F]\cap[D,D])^{\rhd_D\lhd_A})\in S\text{-}ISL(U,A,V,I)$，则$[E,F]\cap[D,D]\leqslant([E,F]\cap[D,D])^{\rhd_D\lhd_A}<X^{\lhd_A}$。我们断言$[E,F]\cap[D,D]<X^{\lhd_D}$，若$[E,F]\cap[D,D]=X^{\lhd_D}=[B,C]\cap[D,D]$，则$X=X^{\lhd_D\rhd_D}=([B,C]\cap[D,D])^{\lhd_D}=([E,F]\cap[D,D])^{\lhd_D}$，矛盾。故$[E,F]\cap[D,D]<X^{\lhd_D}=[B,C]\cap[D,D]$，即$((B\text{-}E)\cap D)\cup((C\text{-}F)\cap D)=D\cap((B\text{-}E)\cup(C\text{-}F))\neq\emptyset$，显然$D\cap DIS_{S\text{-}ISL}((X,[B,C]),(Y,[E,F]))\neq\emptyset$。

(3)$\Rightarrow$(1)。假设$(X,[B,C])=(X,X^{\lhd_A})\in S\text{-}ISL(U,A,V,I)$，$(Y,[E,F])\in PC((X,[B,C]))$，$D\cap DIS_{S\text{-}ISL}((X,[B,C]),(Y,[E,F]))\neq\emptyset$且$X\neq X^{\lhd_D\rhd_D}$，则$X\subset X^{\lhd_D\rhd_D}=X^{\lhd_D\rhd_A}$。由$(X^{\lhd_D\rhd_A},X^{\lhd_D\rhd_A\lhd_A})\in S\text{-}ISL(U,A,V,I)$，可得$(X,X^{\lhd_A})\prec(X^{\lhd_D\rhd_A},X^{\lhd_D\rhd_A\lhd_A})$。我们断言$(X,X^{\lhd_A})\prec(X^{\lhd_D\rhd_A},X^{\lhd_D\rhd_A\lhd_A})$。若存在$(Z,[M,N])\in S\text{-}ISL(U,A,V,I)$，使得$(X,X^{\lhd_A})\prec(Z,[M,N])<(X^{\lhd_D\rhd_A},X^{\lhd_D\rhd_A\lhd_A})$，则$X^{\lhd_D\rhd_A\lhd_A}<[M,N]<X^{\lhd_A}$，故$X^{\lhd_D\rhd_A\lhd_A}\cap[D,D]\leqslant[M,N]\cap[D,D]\leqslant[B,C]\cap[D,D]$。由假设$[M,N]\cap[D,D]<[B,C]\cap[D,D]$，又因为$X^{\lhd_D}\leqslant X^{\lhd_D\rhd_A\lhd_A}$，则$X^{\lhd_D}\leqslant[M,N]\cap[D,D]<[B,C]\cap[D,D]=X^{\lhd_D}$，矛盾，因此，$(X,X^{\lhd_A})\prec(X^{\lhd_D\rhd_A},X^{\lhd_D\rhd_A\lhd_A})$。由假设可知，$D\cap DIS_{S\text{-}ISL}((X,X^{\lhd_A}),(X^{\lhd_D\rhd_A},X^{\lhd_D\rhd_A\lhd_A}))\neq\emptyset$，即$X^{\lhd_D\rhd_A\lhd_A}\cap[D,D]\subset[B,C]\cap[D,D]$。因此$X^{\lhd_D}=X^{\lhd_A}\cap[D,D]\leqslant X^{\lhd_D\rhd_A\lhd_A}\cap[D,D]<X^{\lhd_D}$，矛盾，因此$X^{\lhd_D\rhd_D}=X$，即$D$为格协调集。

定义 11 设(U,A,V,I)为不完备形式背景，S-ISL-差别函数、S-ISM-差别函数、S-ISJ-差别函数、S-ISG-差别函数的定义分别如下：

$$f(\bigwedge_{S\text{-}ISL})=f(\bigwedge_{S\text{-}ISM})=\bigwedge_{H\in \bigwedge_{S\text{-}ISL}}(\bigvee_{h\in H} h)$$

$$f(\bigwedge_{S-ISJ})=\bigwedge_{H\in \bigwedge_{S\text{-}ISJ}}(\bigvee_{h\in H} h)$$

$$f(\bigwedge_{S\text{-}ISG})=\bigwedge_{H\in \bigwedge_{S\text{-}ISG}}(\bigvee_{h\in H} h)$$

由吸收律和分配律，差别函数 f 可被表示成极小析取范式的形式，每一个范式都是(U,A,V,I)的约简。

例 3(续例 2)　不完备形式背景(U,A,V,I) 的 $S\text{-}ISL$-差别矩阵如表 2 所列。

表 2　不完备形式背景(U,A,V,I) 的 $S\text{-}ISL$-差别矩阵

Table 2　The $S\text{-}ISL$-discernibility matrix of incomplete formal context (U,A,V,I)

	$S\text{-}ISC_1$	$S\text{-}ISC_2$	$S\text{-}ISC_3$	$S\text{-}ISC_4$	$S\text{-}ISC_5$	$S\text{-}ISC_6$	$S\text{-}ISC_7$	$S\text{-}ISC_8$	$S\text{-}ISC_9$	$S\text{-}ISC_{10}$	$S\text{-}ISC_{11}$
$S\text{-}ISC_1$											
$S\text{-}ISC_2$	g										
$S\text{-}ISC_3$	ab										
$S\text{-}ISC_4$	d										
$S\text{-}ISC_5$		ab	g								
$S\text{-}ISC_6$			cf								
$S\text{-}ISC_7$			ab								
$S\text{-}ISC_8$				$abeg$	$abdeg$		deg				
$S\text{-}ISC_9$					cf	g					
$S\text{-}ISC_{10}$						abf	cf				
$S\text{-}ISC_{11}$								cf	$abdefg$	deg	

下面基于差别函数计算不完备背景的约简。

$$f(\bigwedge_{S\text{-}ISL})=f(\bigwedge_{S\text{-}ISM})=\bigwedge_{H\in \bigwedge_{S\text{-}ISL}}(\bigvee_{h\in H} h)=g\wedge d\wedge(a\vee b)\wedge(a\vee b\vee e\vee g)\wedge(a\vee b\vee d\vee e\vee g)\wedge(c\vee f)\wedge(d\vee e\vee g)\wedge(a\vee b\vee f)\wedge(a\vee b\vee d\vee e\vee f\vee g)=(a\wedge c\wedge d\wedge g)\vee(a\wedge d\wedge f\wedge g)\vee(b\wedge c\wedge d\wedge g)\vee(b\wedge d\wedge f\wedge g)$$

$$f(\bigwedge_{S\text{-}ISJ})=\bigwedge_{H\in \bigwedge_{S\text{-}ISJ}}(\bigvee_{h\in H} h)=d\wedge(c\vee f)(\wedge(a\vee b\vee d\vee e\vee g)\wedge(a\vee b\vee e\vee g)\wedge(d\vee e\vee g)=(a\wedge c\wedge d\wedge g)\vee(b\wedge c\wedge d\wedge g))\vee(d\wedge f\wedge g)$$

$$f(\bigwedge_{S\text{-}ISG})=\bigwedge_{H\in \bigwedge_{S\text{-}ISG}}(\bigvee_{h\in H} h)=d\wedge(c\vee f)(\wedge(a\vee b\vee d\vee e\vee g)\wedge(a\vee b\vee e\vee g)\wedge(d\vee e\vee g)=(a\wedge c\wedge d\wedge g)\vee(b\wedge c\wedge d\wedge g))\vee(d\wedge f\wedge g)$$

故保持单边区间集概念格结构(交不可约元 c 外延)不变的约简为$\{a,c,d,g\}$，$\{a,d,f,g\}$，$\{b,c,d,g\}$，$\{b,d,f,g\}$；保持并不可约元外延不变的约简为$\{a,c,d,g\}$，$\{a,d,f,g\}$，$\{d,f,g\}$；保持对象单边区间集概念外延不变的约简为$\{a,c,d,g\}$，$\{a,d,f,g\}$，$\{d,f,g\}$。

我们仅给出(U,A_1,V,I_1) 和(U,A_2,V,I_2) 的格图，如图 3、图 4 所示。其中 $A_1=\{d,f,g\}$为保持并不可约元外延不变的约简，$A_2=\{a,c,d,g\}$为保持单边区间集概念格结构(交不可约元外延)不变的约简。

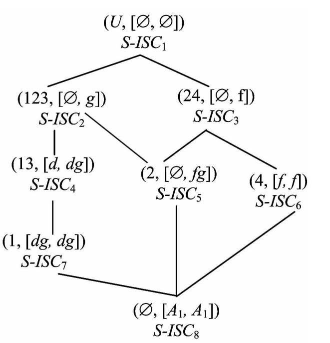

图 3　$S\text{-}ISL(U,A_1,V,I_1)$

Fig. 3　$S\text{-}ISL(U,A_1,V,I_1)$

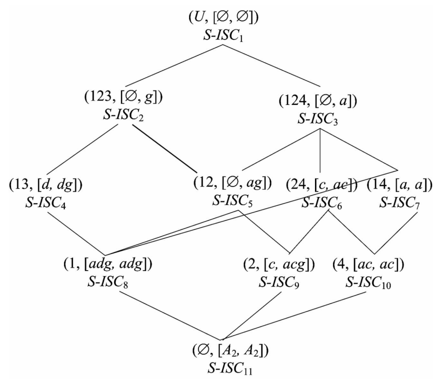

图 4　$S\text{-}ISL(U,A_2,V,I_2)$

Fig. 4　$S\text{-}ISL(U,A_2,V,I_2)$

结束语　本文针对不完备形式背景，提出了 4 种不同的属性约简并研究了它们之间的关系。不同约简反映的不完备形式背景的信息不同，不同情形需要我们甄别对待。由不完备形式背景上的约简定义可知，因为格结构可以反映背景上知识间的承继关系，所以保持单边区间集概念格结构不变的约简可以保持知识间的这种承继关系，但是这种约简需要考虑所有的单边区间集概念，这增加了约简计算过程的复杂度。保持交(并)不可约元外延不变的约简，虽然不能保持知识间的承继关系，但由于不可约元集的稠密性，它在格构建中仍有重要意义，并且在计算过程中仅需要考虑不可约元，这降低了计算的复杂度。由于对象单边区间集概念可被视为信息粒度，因此保持对象单边区间集概念外延不变的约简，在粒计算中有着重要意义，由于它在计算过程中仅需考虑对象单边区间集概念，使得其计算过程最为简单。

针对本文提出的 4 种约简，可根据不同的目的继续定义不同的约简并研究它们之间的关系。

参 考 文 献

[1]　WILLE R. Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts[M]//Riaral I, ed. ordered Sets. Reidel, Dord recht, 1982: 445-470.

[2]　LIU X L, HONG W X, ZHANG T. TCM Differentiation Visua-

lization Methods Based on Formal Concept Analysis[J]. Journal of Yanshan University,2010,34(2):162-168. (in Chinese)
刘旭龙,洪文学,张涛. 基于形式概念分析的中医辨证可视化方法[J]. 燕山大学学报,2010,34(2):162-168.
[3] MISSAOUI R,GODIN R. Extracting Exact and Approximate Rules from Databases [C]// Proceedings of SOFTEKS Workshop on Incompleteness and Uncertainty in Information Systems. 1993:209-222.
[4] GUO X E,WANG J H. Multi-dimensional Concept Lattice and Association Rules Discovery[J]. Journal of Computer Application,2010,30(4):1072-1075. (in Chinese)
郭显娥,王俊红. 多维概念格与关联规则的发现[J]. 计算机应用,2010,30(4):1072-1075.
[5] SUTTON A,MALETIC J I. Recovering UML Class Models from C++:A Detailed Explanation[J]. Information and Software Technology,2007,49(3):212-229.
[6] GODIN R,MISSAOUI R. An Incremental Concept Formation Approach for Learning from Database[J]. Theoretical Computer Science,1994,133(2):387-419.
[7] FREEMAN L,WHITE D. Using Galois Lattices to Represent Network Data[J]. Sociological Methodology,1993,23:127-146.
[8] HO T B. An Approach to Concept Formation Based on Formal Concept Analysis[J]. IEICE Transactions on Information and Systems,1995,E78-D(5):553-559.
[9] PAWLAK Z. Theoretical Aspects of Reasoning about Data[M]. Boston,Kluwer Academic Publishers,1991.
[10] WEI L. The Theory and Methods of Rough Sets and Concept Lattice Reduction[D]. Xi'an:Xi'an Jiaotong University,2005. (in Chinese)
魏玲. 粗糙集与概念格约简理论与方法[D]. 西安:西安交通大学,2005.
[11] NOURINE L,RAYNAUD O. A Fast Algorithm for Building Lattices[J]. Information Processing Letters,1999,71:199-204.
[12] ZHANG W X,QIU G F. Uncertain Decision Making Based on Rough Sets[M]. Beijing:Tsinghua University Press,2005:185-193. (in Chinese)
张文修,仇国芳. 基于粗糙集的不确定决策[M]. 北京:清华大学出版社,2005:185-193.
[13] WEI L,WAN Q,QIAN T,et al. An Overview of Triadic Concept Analysis[J]. Journal of Northwest University (Natural Science Edition),2014,44(5):689-699. (in Chinese)
魏玲,万青,钱婷,等. 三元概念分析综述[J]. 西北大学学报(自然科学版),2014,44(5):689-699.
[14] REN R S,WEI L. The attribute reductions of three-way concept lattices[J]. Knowledge-Based Systems,2016,99(C):92-102.
[15] BURMEISTER P,HOLZER R. On the treatment of incomplete knowledge in formal concept analysis[C]// Conceptual Structures:Logical,Linguist,and Computa-tional Issues. Berlin,Heidelberg:Springer-Verlag,2000:385-398.
[16] LI J H,MEI C L,LV Y J. Incomplete decision contexts:Approximate concept construction,rule acquisition and knowledge reduction[J]. International Journal of Approximate Reasoning,2013,54(1):149-165.
[17] YAO Y. Interval sets and three-way concept analysis in incomplete contexts[J]. International Journal of Machine Learning & Cybernetics,2017,8(1):3-20.
[18] LI M Z,WANG G Y. Approximate concept construction with three-way decisions and attribute reduction in incomplete contexts[J]. Knowledge-Based Systems,2016,91:165-178.
[19] DAVEY B A,PRIESTLEY H A. Introduction of Lattice and Order[M] // Cambridge University Press. Cambridge United Kingdom,1990.
[20] ZHANG W X,WEI L,QI J J. Attribute reduction theory and approach to concept lattice[J]. Science in China Series,2005,48(6):713-726.

(上接第61页)

[8] WANG W W,LI X P,FENG X C. Summarization of sparse subspaces[J]. Journal of Acta Automatica Sinica,2015,41(8):1373-1384. (in Chinese)
王卫卫,李小平,冯象初. 稀疏子空间聚类综述[J]. 自动化学报,2015,41(8):1373-1384.
[9] WANG S,LUO L,ZHANG Z. SPSD matrix approximation vis column selection:theories,algorithms,and extensions [J]. Journal of Machine Learning Research,2016,17(5):1-49.
[10] WANG S,ZHANG Z,ZHANG T. Towards more efficient SPSD matrix approximation and CUR matrix decomposition [J]. Journal of Machine Learning Research,2016,17(210):1-49.
[11] ZHU L,LEI J S,BI Z Q. A soft subspace clustering algorithm based on data stream[J]. Journal of Software,2013,24(11):2610-2627. (in Chinese)
朱林,雷景生,毕忠勤. 一种基于数据流的软子空间聚类算法[J]. 软件学报,2013,24(11):2610-2627.
[12] BACH F R,JORDAN M I. Learning spectral clustering [J]. Advances in Neural Information Processing Systems,2004,16(2):2006.
[13] DEMPSTER A. Maximum likelihood from incomplete data via the EM algorithm [J]. Journal of the Royal Statistical Society,1977,39(1):1-38.
[14] SPIELMAN D A. Spectral graph theory and its applications [C]// IEEE Symposium on Foundations of Computer Science. 2007:29-38.
[15] CAI X Y,DAI G Z,YANG L B. A review of spectral clustering algorithm [J]. Journal of Computer Science,2008,35(7):14-18. (in Chinese)
蔡晓妍,戴冠中,杨黎斌. 谱聚类算法综述[J]. 计算机科学,2008,35(7):14-18.
[16] ALZATE C,SUYKENS J A K. Multiway Spectral Clustering with Out-of-Sample Extensions through Weighted Kernel PCA [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2010,32(2):335-347.
[17] LAUER F,SCHNOOR C. Spectral clustering of linear subspaces for motion segmentation [J]. IEEE International Conference on Computer Vision,2009,30(2):678-685.
[18] JIA H,DING S,DU M. Approximate normalized cuts without eigen-decomposition [J]. Journal of Information Sciences,2016,374(5):135-150.
[19] SHI J,MALIK J. Normalised cuts and image segmenerarion[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2000,22(8):888-905.
[20] DING C,LI T,JORDAN M I. Convex and semi-nonnegative matrix factorizations[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2010,32(1):45-55.