УДК 519.8

ПСЕВДОБУЛЕВЫ КАНОНИЧЕСКИЕ ОПТИМИЗАЦИОННЫЕ МОДЕЛИ И МАТРОИДЫ

В.И. Донской

Таврический национальный университет им. В.И. Вернадского
факультет математики и информатики
пр-т Вернадского, 4, г. Симферополь, 95007, Украина
e-mail: *donskoy@ccssu.crimea.ua*

Abstract

The accuracy of decision making based on the models of optimal choice with disjunctive constraint: $\max f(\hat{x}) = \sum_{i=1}^{n} \omega(x_i)$ with the condition $D(\hat{x}) = \bigvee_{j=1}^{m} K_j(\hat{x}) = 1$ for the case when the condition is exact, but the information for the goal function is presented only by the weight order $\omega(x_1) \geqslant \dots\ \omega(x_n) \geqslant 0$, is estimated.

Введение

Изучение и разработка моделей выбора оптимальных решений на основе накопленных знаний - важное направление в области интеллектуального анализа н обработки информации. И настоящее время вопросы извлечения и представления знаний изучены гораздо глубже, чем подходы к построению процедур выбора оптимальных (с точностью, определяемой имеющейся неполной начальной информацией) решений на основе полученных знаний [1],[2],[3]. Выбор решений в интеллектуализированных компьютерных системах должен осуществляться на основе синтеза моделей, адекватных имеющейся (но, как правило, неполной) начальной информации, представленной знаниями и прецедентами - фактами.

Классический подход к синтезу моделей принятия решений предполагает определение области допустимых решений и оценочной (скалярной или векторной) целевой функции. Для выбора моделей в рамках классического подхода используется информация двух видов: качественная информация о классе, в котором отыскивается модель, и информация о параметрах модели. Основным недостатком классических моделей выбора решений является их «условная» согласованность с моделируемыми объектами и явлениями. При выборе классических моделей заранее предполагается выполнение некоторых, обычно жестких, условий, определяющих вид модели: *независимо от имеющихся эмпирических данных модель фиксируется*, и далее требуется получить значения всех ее числовых параметров.

Неклассическое индуктивное моделирование основывается на прецедентах (фактах) и знаниях, представляющих собой, главным образом, отношения определенного вида, которым удовлетворяют или априорно должны удовлетворять прецеденты.

Использование псевдобулевых моделей, о которых будет идти речь в статье, предполагает возможным предикатное описание объектов, т. е. использование свойств-предикатов в качестве семантической базы представления проблемной области. Предикатам ставятся в соответствие булевы переменные, на основе которых могут строиться модели ограничений и целевых функций. Задание свойств-предикатов обычно является основой определения элементов баз знаний продукционного типа, накопления наборов эмпирических значений в базах экспериментальных данных, извлечения логических закономерностей - аналогии, представительных наборов, деревьев решений. На базе предикатного описания объектов сроятся дизъюнктивные нормальные формы (ДНФ) - логические описания областей индуктивной и дедуктивной выводимости, выводимости по аналогии [5], [7]. Изложенное выше обосновывает *необходимость дальнейшего исследования псевдобулевых канонических оптимизационных моделей, получаемых в результате неклассического, в частности, индуктивного моделирования, и постановки следующих задач. Представляет интерес выяснение условий, при которых эти модели определяют области допустимых решений, удовлетворяющие свойствам матроидов, нахождение и обоснование методов поиска приближенных решений в случаях неполной начальной информации.*

Целью данной работы является *оценка точности решений*, выбираемых *на основе модели*, которая синтезируется по прецедентной информации и/или знаниям продукционного типа в форме *с дизъюнктивным ограничением*. Замечательным свойством такой модели, как установлено ниже, является то, что *информации только о порядке слагаемых аддитивной целевой функции с неотрицательными весами оказывается достаточным, чтобы, имея ДНФ ограничение, оценить точность выбранного* $GREEDY$ *алгоритмом решения.*

1. Синтез канонических псевдобулевых моделей и их свойства

В работе [4] показано, что любая задача псевдобулевой оптимизации вида

$$extr\ f(\tilde{x}), \quad \tilde{x} \in W \subseteq B^n = \{0,1\}^n,$$

где $f\,:\,B^n \to \mathbb{R}$, может быть представлена в канонической форме с дизъюнктивным ограничением

$$\begin{cases} extr\, f(\tilde{x}), \\ D(\tilde{x}) = 1, \end{cases} \tag{1}$$

где $D(\tilde{x}) = \bigvee_{j=1}^{m} x_{j_1}^{\sigma_{j_1}} \cdot \ldots \cdot x_{r_j}^{\sigma_{j r_j}}$ - произвольная ДНФ характеристической функции φ_W множества W, определяемая выражением $\varphi_W(\tilde{x}) = \begin{cases} 1, & \tilde{x} \in W, \\ 0, & \tilde{x} \notin W. \end{cases}$ В данной статье рассматривается важный для приложений случай, когда целевая функция является аддитивной с неотрицательными весами: $f(\hat{x}) = \sum_{i=1}^{n} \omega(x_i)$, $\omega(x_i) \geqslant 0$.

Далее класс псевдобулевых функций обозначается $PS_2(n) = \{f\,:\,B^n \to \mathbb{R}\}$.

Значение канонической формы (1) становится ясным с учетом следующих полученных в [4] результатов.

- Как уже отмечено выше, *любая задача псевдобулевой оптимизации представима в форме* (1) с ДНФ ограничением.
- *Если задача оптимизации линейной псевдобулевой функции с ограничениями-неравенствами приводится к эквивалентной форме с ДНФ ограничением за число шагов, ограниченное полиномом от размерности задачи, то она разрешима за полиномиальное время.*
- *Для любой нелинейной задачи безусловной оптимизации функции* $f \in PS_2(n)$ *найдется эквивалентная ей задача оптимизации линейной функции* $g \in PS_2(m)$ *в форме с дизъюнктивным условием, где* m *- число слагаемых приведенного полинома для функции* f *без учета свободного члена.*

Задачи псевдобулевой оптимизации, представленные в форме с ограничениями типа неравенств или в ином виде, как следует из указанных ниже фактов, совсем не обязательно специально приводить к форме с ДНФ ограничением, поскольку такое приведение равносильно решению задачи и в некоторых случаях может оказаться сложнее, чем решение другим методом. Важность канонической формы объясняется прежде всего тем, что уже *изучены методы построения ДНФ ограничений, использующие о качестве начальной информации прецеденты, накопленные в базах эмпирических данных и/или продукции, составляющие базы знаний* [5], [7]. Оказалось, что используя указанную начальную информацию, построить ограничения в форме ДНФ гораздо проще (и точнее), чем в другой форме (например, н виде неравенств).

2. Матроиды и канонические модели

Пусть $E = \{e_1, \ldots, e_n\}$ - конечное множество, $B(E)$ - булеан над E, $J \subseteq B(E)$ - зафиксированное семейство подмножеств. Семейство J называют *системой независимости, если*

$$\emptyset \in J; \quad (C \in J) \wedge (A \subset C) \Rightarrow A \in J \tag{2}$$

Множества, входящие в J, называются *независимыми.*

Пару $\langle E, J \rangle$ называют *матроидом*, . если непустое семейство J является системой независимости, и все максимальные по включению независимые подмножества (*базы*) любого множества $D \subseteq E$ равномощны. Мощность баз множества D называется его *рангом* $r(D)$, а ранг $r(E)$ множества E называется *рангом матроида.* Любые две базы матроида - максимальные независимые множества в E - равномощны и имеют ранг $r(E)$ [8].

Если в задаче

$$\begin{cases} \max_{\{S : S \in J\}} \sum_{e \in S} \omega(e), \\ J \subseteq B(E), \\ \omega \ : \ E \to R^+ \end{cases} \tag{3}$$

пара $\langle E, J\rangle$ является матроидом, то, согласно теореме Радо-Эдмондса [8], жадный $(GREEDY)$ алгоритм обеспечивает точное решение задачи (3), притом с полиномиально ограниченной трудоемкостью, если только свойство $\ll S \in J \gg$ для любого $S \subseteq E$ вычислимо с полиномиальной сложностью.

Форма представления задачи (3) является обобщенной постановкой многих задач дискретной оптимизации. Входящие в нее множество E и семейство $J \subseteq B(E)$ образуют пару $\langle E, J\rangle$, которая, возможно, не является матроидом, и в таких случаях $GREEDY$ алгоритм не гарантирует нахождения точного решения.

Свяжем булевы переменные $x_1, \ldots, x_n$ из задачи (1) с элементами множества $E_{\tilde{x}} = \{e_1, \ldots, e_n\}$ так, что вектор $\tilde{\alpha} = \alpha_1, \ldots, \alpha_i, \ldots, \alpha_n$ 0-1-значений переменных $x_1, \ldots, x_n$ будет определять множество $S_{\tilde{\alpha}} = \{e_i \quad : \quad \alpha_i = 1\} \subseteq E$. Вектор $\tilde{0} = (0, 0, \ldots, 0)$ соответствует пустому множеству. Тогда, если $D(\tilde{x}) = \bigvee_{j=1}^{m} K_j(\tilde{x})$ - ДНФ, определяющая ограничение канонической модели (1), то условие (уравнение) $D(\tilde{x}) = 1$ равносильно условию $S_{\tilde{x}} \in J$ (соотношению $\bigcup_{\tilde{x} \in \{\tilde{x}: D(\tilde{x})=1\}} S_{\tilde{x}} = J$). Здесь J - некоторое семейство множеств, которое может быть или не быть системой независимости. Это семейство J определяется ДНФ $D(\tilde{x})$ однозначно: $J = J(D(\tilde{x}))$. Поэтому далее вместо семейства $J(D(\tilde{x}))$ будет использоваться определяющим его ДНФ $D(\tilde{x})$.

Очевидно, выполняется соотношение $J(D(\tilde{x})) \subseteq B(E_{\tilde{x}})$, где $B(E_{\tilde{x}})$ - булеан. Важно установить, является ли пара $\langle E_{\tilde{x}}, D(\tilde{x})\rangle$ матроидом или близким по свойствам к матроиду объектом.

Теорема 1. *Если в канонической модели (1) пара $\langle E_{\tilde{x}}, D(\tilde{x})\rangle$ является матроидом, то выполняются условия:*

- 1^0 $D(\tilde{0}) = 1$;
- 2^0 *Любая простая импликата булевой функции f_D, определяемой ДНФ $D(\tilde{x})$, не содержит положительных литералов и имеет ранг $r_{\min} = n - r(M)$, где $r(M)$ - ранг матроида $\langle E_{\tilde{x}}, D(\tilde{x})\rangle$.*

Доказательство. Необходимость 1^0 немедленно следует из (2) в определении системы независимости, которая должна содержать в себе пустое множество.

Пусть f_D - булева функция, определяемая ДНФ $D(\tilde{x})$, и K - простая импликата для f_D. Если бы K содержала положительный литерал, то она имела бы вид $K = x_s H$. Но тогда, согласно определению матроида, импликатой для f_D была бы и конъюнкция $K_1 = \bar{x}_s H$. Поскольку $\bar{x}_s H \vee x_s H = H$, то это противоречило бы тому, что K - простая импликата. Следовательно, в K содержится некоторое количество r_k только отрицательных литералов, из которых ни один удалить нельзя, поскольку K - простая импликата.

Если $K = \bar{x_{j_1}} \wedge \ldots \wedge \bar{x_{j_{r_k}}}$, то $\tilde{\sigma} \ : \ \sigma_i = \begin{cases} 0, & i \in \{j_1, \ldots, j_{r_k}\}, \\ 1, & i \notin \{j_1, \ldots, j_{r_k}\}, \end{cases}$ - база матроида M. Все базы матроида равномощны, откуда $r_k = n - r(M)$. Теорема доказана.

Следующий пример позволяет убедиться, что условия теоремы 1 не являются достаточными для того, чтобы пара $\langle E_{\tilde{x}}, D(\tilde{x})\rangle$ была матроидом. В задаче

$$\max(5x_1 + 4x_2 + 4x_3 + x_4)/\bar{x}_2\bar{x}_3 \vee \bar{x}_1\bar{x}_4 = 1$$

$GREEDY$ алгоритм выберет допустимое решение $(1, 0, 0, 1)$, имеющее вес 6, в то время как решение $(0, 1, 1, 0)$ имеет больший вес, равный 8. Дизъюнктивное ограничение $\bar{x}_2\bar{x}_3 \vee \bar{x}_1\bar{x}_4 = 1$ удовлетворяет условию теоремы 1, но соответствующая пара $\langle E_{\tilde{x}}, D(\tilde{x})\rangle$ не может быть матроидом, поскольку это противоречило бы теореме Радо-Эдмондса.

Таким образом. если ДНФ $D(\tilde{x})$ в задаче (1) в канонической форме с ограничением $D(\tilde{x}) = 1$ можно преобразовать в ДНФ $D^{-}(\tilde{x})$, не имеющую положительных литералов и реализующую ту же функцию, что и $D(\tilde{x})$, то ограничение $D(\tilde{x}) = 1$ определяет некоторую систему независимости $J \subset B^n$. Если при этом в ДНФ $D(\tilde{x}) = 1$ все конъюнкции являются простыми импликатами и имеют одинаковый ранг, то выполняются необходимые, но, вообще говоря, не достаточные условия того, что $\langle E_{\tilde{x}}, D(\tilde{x})\rangle$ - матроид.

Рис.1 иллюстрирует матроид M, все базы которого составляют $r(M)$ - слой куба B^n, то есть соответствуют наборам его вершин, каждая из которых содержит ровно $r(M)$ единиц.

3. Неполнота начальной информации и обоснование решений

В рамках настоящей работы неполная начальная информации в задаче выбора решения описывается следующим обратом: целевая функция достоверно является аддитивной, но не задана точно; имеется лишь информация о «порядке» весов

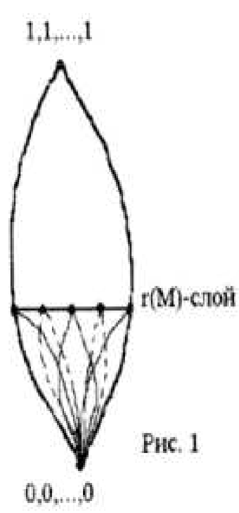

Рис. 1

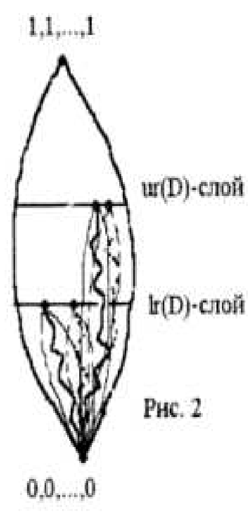

Рис. 2

переменных: $\omega(x_{i_1}) \geqslant \ldots\omega(x_{i_n}) \geqslant 0$. Полагается, что начальная информация для выбора решения порождена (не известной полностью) моделью $\max f(\tilde{x}) = \sum_{i=1}^{n}\omega(x_i)$ при условии $D(\tilde{x}) = 1$ - ДНФ, реализуюшая характеристическую функцию множества допустимых решений задачи. Далее будем налагать, что ДНФ $D(\tilde{x})$ задана (получена некоторым способом [5], [7]) и точно описывает ограничения. Будем обозначать $\tilde{x}_G$ и $f(\tilde{x}_G)$ соответственно точку, соответствующую решению этой задачи, и значение целевой функции и этой точке, найденные при помощи $GREEDY$ алгоритма, а $\tilde{x}_{opt}$ и $f(\tilde{x}_{opt})$ - точку и значение целевой функции существующего, но неизвестного оптимального решения.

Пусть J - семейство независимости, но пара $\langle E, J\rangle$, вообще говоря, может не быть матроидом. Обозначим для любого множества $F \subseteq E$

$$lr(F,J) = \min\{|C| \,:\, C \,-\, \text{база в } F\}$$

$$ur(F,J) = \max\{|C| \,:\, C \,-\, \text{база в } F\}$$

Число

$$k(J) = \min_{F\subseteq E:\, ur(F,J)\neq 0} \frac{lr(F,J)}{ur(F,J)}$$

называется *кривизной системы независимости* J и удовлетворяет двойному неравенству $0 < k(J) \leqslant 1$, причем $k(J) = 1$ только тогда, когда $\langle E, J\rangle$ - матроид [9]. Пусть S_G - множество, отобранное $GREEDY$ алгоритмом решения задачи (3), а S_0 - множество, на котором в действительности достигается искомый экстремум.

Доказано [9], что для любой неотрицательной функции ω имеет место неравенство $\frac{\omega(S_G)}{\omega(S_0)} \geqslant k(J)$. Далее используется аналогичный подход па основе нового понятия *кривизны* ДНФ, описывающей систему независимости.

Определение 1. Кривизной ДНФ $D = \bigvee_{j=1}^{m} K_j$ называется число $k(D) = \frac{\min_{1\leqslant j\leqslant m}\{n-r_j^-\}}{\max_{1\leqslant j\leqslant m}\{n-r_j^-\}}$, где r_j^- - число отрицательных литералов в конъюнкции K_j.

Будем обозначать $lr(D) = \min_{1\leqslant j\leqslant m}\{n - r_j^-\}$ и $ur(D) = \max_{1\leqslant j\leqslant m}\{n - r_j^-\}$.

Определение 2. Пусть f_D - булева функция, определяемая ДНФ D. $S(f_D)$ - множество тех простых импликат функции f_D, r_L^- - число отрицательных литералов в простой импликате $L \in S(f_D)$. *Кривизной булевой функции* f_D называется число $k(f_D) = \frac{\min_{L\in S(f_D)}\{n-r_L^-\}}{\max_{L\in S(f_D)}\{n-r_L^-\}}$

Будем обозначать $lr(f_D) = \min_{L\in S(f_D)}\{n - r_L^-\}$ и $ur(f_D) = \max_{L\in S(f_D)}\{n - r_L^-\}$.

Теорема 2. *Пусть в задаче в канонической форме с информацией о порядке весов переменных* $\omega(x_{i_1}) \geqslant \ldots\omega(x_{i_n}) \geqslant 0$ *и точно заданным ограничением* $\bigvee_{j=1}^{m} K_j = 1$ *кривизна ДНФ* $k(D) = \frac{\mu}{M}$, *где* $\mu = \min_{1\leqslant j\leqslant m}\{n - r_j^-\}$, $M = \max_{1\leqslant j\leqslant m}\{n - r_j^-\}$, *и* *первые μ переменные $x_{i_1}, \ldots, x_{i_\mu}$, согласно упорядочению $\omega(x_{i_1}) \geqslant \ldots \omega(x_{i_n}) \geqslant 0$, не входят с инверсиями хотя бы в одну из конъюнкций $K_1, \ldots, K_i, \ldots, K_m$. Тогда $\frac{f(\tilde{x}_G)}{f(\tilde{x}_{opt})} \geqslant k(D)$.*

Доказательство. Заметим, что $1 \leqslant \mu \leqslant M < n$. Число элементов (единиц в булевом векторе $\tilde{x}$) в любом выбранном $GREEDY$ алгоритмом множестве лежит в отрезке $[\mu, M]$. Убедимся в том, что $f(\tilde{x}_G) \geqslant k(D)f(\tilde{x}_{opt})$. Пуль Ω - допустимое множество с максимальным весом, $|\Omega| \leqslant M$, $f(\tilde{x}_{opt}) = f(\Omega) = \sum_{e \in \Omega} \omega(e)$, и $f(\tilde{x}_G) = f(G) = \sum_{e \in \Omega} \omega(e)$, где G - множество элементов, отобранных $GREEDY$ алгоритмом. Очевидно, $|G| \geqslant \mu$. Тогда

$$k(D)f(\tilde{x}_{opt}) = kf(\Omega) = \frac{\mu}{M} f(\Omega) = \mu \frac{1}{M} \sum_{e \in \Omega} \omega(e) \leqslant \mu \frac{1}{|\Omega|} \sum_{e \in \Omega} \omega(e) =$$

$$= \mu a_\Omega \leqslant \mu \frac{1}{\mu} \sum_{e \in \Omega} \omega(e) \leqslant \sum_{e \in \Omega} \omega(e) = f(G) = f(\tilde{x}_G).$$

Здесь a_Ω - средний вес элементов множества Ω, а Ω_μ - множество, состоящее из первых μ по порядку (больших) элементов множества Ω. Приведенная цепочка неравенств доказывает теорему.

В силу того, что ДНФ, представляющая произвольную характеристическую функцию ограничений, не единственна, значение кривизны ДНФ будет зависеть от входящих в нее конъюнкций.

На рис.2 булевы наборы $ur(D)$ - слоя имеют норму M, а наборы $lr(D)$ - слоя - норму μ.

Процесс выбора и оценивания решений состоит в следующем.

1) Получить каноническую форму задачи с ограничением, определяемым ДНФ D (при этом полагается заданным упорядочение переменных по весу).
2) Упростить ДНФ D, стараясь получить минимальную или тупиковую ДНФ D^*.
3) Вычислить кривизну ДНФ $k(D^*)$ и проверить выполнение условия теоремы 2
4) Выбрать в D^* конъюнкцию K_{j^*} следуя правилу $GREEDY$ алгоритма.
5) Внять в качестве решения булев вектор, обращающий в единицу конъюнкцию K_{j^*} и имеющий единичные значения всех переменных, не входящих в K_{j^*} Если ДНФ D точно определяет область допустимых решений, то $f(\tilde{x}_G) \geqslant k(D)f(\tilde{x}_{opt})$.
6) Если для заданной точности ε выполнено $k(D^*) \geqslant 1 - \varepsilon$ и выполнено условие теоремы 2, то имеется гарантия, что решение, найденное $GREEDY$ алгоритмом будет отличаться от оптимального решения $\tilde{x}_{opt}$ не более чем на $\varepsilon f(\tilde{x}_{opt})$ или не более чем на $100\varepsilon\%$.

Заключение

Основными результатами данной статьи являются теоремы 1 и 2, в которых получены необходимые условия соответствия канонической модели с ДНФ ограничением структуре матроида и оценка точности решения задачи, представленной канонической моделью, при использовании $GREEDY$ алгоритма.

Главный вывод, который следует из полученных в статье результатов, состоит в следующем. По виду ДНФ канонической формы, представляющей ограничения в задаче псевдобулевой оптимизации, можно судить, насколько целесообразно использование $GREEDY$ алгоритма для ее решения в случае, когда начальная информация об аддитивной целевой функции представлена только порядком весов переменных.

В дальнейших работах было бы интересно изучить соотношения между величинами $k(D)$ и $k(f_D)$ для произвольных ДНФ D, описывающих ограничения в канонической модели, и получить оценки точности при более слабых условиях, чем использованные в теореме 2.

Список литературы

1. *Hand D.J., Mannila H.. Smyth P.* Principles of Data Mining. - MIT Press, 2001. - 425 p.
2. *Han J., Kamber M.* Data Mining: Concepts and Techniques. - Morgan Kaufmanti, 2000. - 550 p.
3. */ Fayyad V., Piatetsky-Shapiro G., Smyth P.* From Data Mining to Knowledge Discovery in Databases // AI Magazine. - 1996. -17. - P.37-54.
4. *Донской В.И.* Задачи псевдобулевой оптимизация с дизъюнктивным ограничением // Журнал вычислительной математики и мат.физики. - 1994. - Т.34, №3. - С.389-398.
5. *Донской В.И.* Логические продукционные системы: аиализ и синтез // Кибернетика и системный анализ. - 1994. - №4. - С.11-22.
6. *Донской В.И.* Слабоопределенные задачи линейного булевого программирования //Журнал вычислительной математики и мат.физики. 1988- - Х28, №9. - С. 1379-1385.
7. *Donskoy V.* Pseudo-Boolean scalar optimization models with incomplete information, GMOOK Newsletter. - 1996. - №1/2. - P.20-26.
8. *Edmonds J.* Matroids and Greedy Algorithms // Math Programming . 1971. - P.127-136.
9. *Hausmann D., Korte B.* Lower bounds on the worst-case complexity of some oracle algorithms // Discrete Math. - 1978. - Vol.24, №3.- P.85-120