УДК 334.02, 519.08

ПРОБЛЕМЫ СИНТЕЗА МЕХАНИЗМА КОМПЛЕКСНОГО ОЦЕНИВАНИЯ НА ОСНОВЕ ОБУЧАЮЩЕГО НАБОРА ДАННЫХ

В.Н. Бурков
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: vlab17@bk.ru

Н.А. Коргин
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: nkorgin@ipu.ru

О.Л. Марин
ООО «ПК Аквариус»
Россия, 108811, Москва, Киевское ш. 6, стр.1
E-mail: 2marin.oleg@gmail.com

Ключевые слова: механизм комплексного оценивания, идентификация.

Аннотация: Механизм комплексного оценивания позволяет решить задачу оценки и ранжирования объектов, по набору их оценок по многим критериям на основе агрегирования оценок по разным критериям в структуру два-дерева с помощью матриц свертки в узлах дерева. В докладе формулируется проблема реализуемости набора обучающего набора данных с помощью механизма комплексного оценивания и приводится итеративный алгоритм проверки реализуемости.

1. Введение

Механизмы комплексного оценивания на протяжении уже достаточно длительного времени применяются в практике управления социально-экономическими системами, см., например, [1]. Однако, вплоть до настоящего момента основой подходов к синтезу структуры два дерева и матриц свертки в его узлах осуществлялся на основе оценки сравнительной важности критериев, являющихся листьями дерева комплексного оценивания, см. [2, 3]. Подходам, позволяющим формализовать процесс синтеза КО на основе наборов обучающих данных, внимания уделено не было. В то время как в смежных областях теории принятия решений, использующих древовидные структуры для синтеза правил принятия решений, данной проблематике уделяется достаточно много внимания, в частности, в области деревьев принятия решений, см., например, [4]. Однако принципиальные различия в подходах к применению деревьев (в т.ч., двоичных) по построении деревьев поддержки принятия решений и механизмов комплексного оценивания не позволяют применить полученные результаты напрямую.

В данном докладе исследуется вопрос – возможно ли реализовать набор обучающих данных с помощью какого-либо механизма комплексного оценивания – приводится ряд формальных постановок данной задачи, примеры, иллюстрирующие актуальность данных постановок и приводится итеративный алгоритм, позволяющий решить ряд поставленных задач.

2. Основные определения и постановка задачи

Рассмотрим конечное множество критериев $M=\{1,...,m\}$, по которым производится оценка некоторого объекта(ов). Для каждого критерия $i \in M$ задано допустимое множество оценок K_i, из которого для объекта определяется оценка по данному критерию $k_i \in K_i$, $k=(k_1,...,k_m)$ - вектор оценок объекта по всем критериям. По всей совокупности значений критериев для объекта из множества допустимых значений K_M необходимо определить «интегральную оценку» объекта - $k_M \in K_M$.

Определение 1. Механизм комплексного оценивания на основе два-дерева – это отображение $KO_M(\cdot):\prod_{i\in M} K_i \to K_M$, в котором над множеством критериев M строится два-дерево – ориентированный граф $G=(V,E)$:

1. $V=M \cup \tilde{M}$, $\tilde{M}=\{m+1,2m-1\}$,
2. $E=\{e_{ij}\} \subseteq V \times V$,
   a. $\forall i \in V \setminus \{2m-1\}\ \exists!\, j \in \tilde{M} \setminus \{i\}:\ e_{ij}=1$, а $\forall t \in V \setminus j\ e_{it}=0$,
   b. $\forall j \in M\ \forall i \in V\ e_{ij}=0$,
   c. $\forall j \in \tilde{M}\ \exists!\{l,r\} \in V \setminus \{j\} \times V \setminus \{j\}: e_{lj}=1, e_{rj}=1$.

Для каждого промежуточного узла дерева $j \in \tilde{M}$ определяется множество допустимых значений K_j оценки в данном узле $k_j \in K_j$, $K_{2m-1}=K_M$ и отображение $f_i(\cdot):K_l \times K_r \to K_j$, где $\{l,r\} \in V \setminus \{j\} \times V \setminus \{j\}: e_{lj}=1, e_{rj}=1$.

Для КО на основе *матричных сверток* множество допустимых значений в узлах дерева и интегральной оценки представляется дискретной (или лингвистической шкалой), т.е. $\forall j \in \tilde{M}\ K_j \subseteq N$. При этом для исходных критериев множества допустимых значений оценок либо изначально берутся также дискретными, либо задаются функции пересчета из исходных значений оценок по критериям в некоторые дискретные шкалы – т.е. $\forall j \in M\ K_j \subseteq N$. При этом, в большинстве постановок подразумевается, что шкала для исходных критериев, узлов дерева свертки и интегральной оценки одна и та же - $\forall j \in M \cup \tilde{M}\ K_j = K_M \subseteq N$. А отображения $f_i(\cdot):K_l \times K_r \to K_i$ определяются *матрицами свертки*, в которых строка матрицы определяется значением критерия для одного из сворачиваемых наборов, столбец – значением для второго, а значение в ячейке определяет значение критерия двух сворачиваемых наборов.

Для удобства дальнейших выкладок приведем определение механизм комплексного оценивания на основе два-дерева и матричных сверток на основе дихотомического итеративного разбиения всего множества критериев:

Определение 2. Механизм комплексного оценивания на основе два-дерева и матричных сверток $KO_M(\cdot) = (G_M, WG_M(\cdot)) : \prod_{i \in M} K_i \to K_M$ задается набором подмножеств критериев $G_M = \{S_t\}_{t \in \{1,\dots,2m-2\}}$:

$$\forall S \in G_M : S \neq \varnothing,\ S \subset M;$$
$$\forall i \in M\ \exists! S \in G_M : S = \{i\};$$
$$\forall S \in G_M :$$
$$1.\ \exists! \{\tilde{S}, \hat{S}\} \subseteq G_M \setminus S \cup M : S = \hat{S} \setminus \tilde{S};$$
$$2.\ \text{если } \#S \geq 2\ \ \exists! \{L, R\} \subseteq G_M \setminus S : L = S \setminus R;$$

и набором матричных сверток $WG_M(\cdot) = \{W_{S,L}(\cdot)\}_{\forall \{S,L\} \subseteq G_M \cup M : L \subset S, S \setminus L \subset G_M}$, где $W_{S,L}(\cdot) : K_L \times K_{S \setminus L} \to K_S$, $S \subseteq M, L \subset S$.

При этом, в большинстве постановок подразумевается, что шкала для исходных критериев, узлов дерева свертки и интегральной оценки одна и та же – $\forall W_{S,L}(\cdot) \in WG_M(\cdot)\ \ W_{S,L}(\cdot) : K_M \times K_M \to K_M$.

Перейдем к формальной постановке задачи синтез механизма комплексного оценивания на основе обучающего набора данных. Обозначим $q = (k, k_M)$ - вектор оценок по всем критериям и его интегральную оценку, $Q \subset K \otimes K_M$, $K = \prod_{i \in M} K_i$ - обучающее множество векторов или обучающий набор. В дальнейшем будем рассматривать *непротиворечивые множества* – $\forall \{q, \tilde{q}\} \subseteq Q\ k \neq \tilde{k}$. Т.е. в обучающем наборе не должно быть векторов, в которых значения оценок по всем критериям совпадают, а интегральные оценки – нет. Обучающий набор будем называть *полным*, если $\forall k \in K\ \exists q \in Q$: $q = (k, k_M)$ в противном случае будем называть его *неполным*. Для неполного Q будем обозначать $\hat{Q}$ произвольный содержащие его полный обучающий набор – $Q \subset \hat{Q}$.

Введем отношение доминирования векторов оценок по критериям следующим образом – $k \succ \tilde{k}, \{k, \tilde{k}\} \subseteq K$ если $\forall i \in M\ \ \tilde{k}_i \leq k_i$.

Определение 3. Обучающий набор Q является монотонным, если $\forall \{q, \tilde{q}\} \subseteq Q$ верно, что если $k \succ \tilde{k}$, то $\tilde{k}_M \leq k_M$. При этом $q \succ \tilde{q}$.

Будем говорить, что некоторое отображение $f_Q(\cdot) : K \to K_M$ реализует обучающий набор $Q \subset K \otimes K_M$, если $\forall q \in Q\ \ f_Q(k) = k_M$.

Для большинства задач комплексного оценивания, для которых разрабатывались и применялись механизмы комплексного оценивания подразумевают *монотонность* матричных сверток.

Определение 4. Матричная свертка $W_{S,L}(\cdot) : K_L \times K_{S \setminus L} \to K_S$ монотонна, если

$$\forall k, \tilde{k}, L, S : L \subset S \subseteq M :$$
$$\{k_L, k_{S \setminus L}\} \succ \{\tilde{k}_L, \tilde{k}_{S \setminus L}\} \Rightarrow W(k_L, k_{S \setminus L}) \geq W(\tilde{k}_L, \tilde{k}_{S \setminus L}).$$

На основе введенных определений можно сформулировать следующие постановки задачи синтеза механизма комплексного оценивания на основе набора обучающих данных.

Задача 1. (идентификация КО с единой шкалой). Пусть $m > 2$, $K_M \in N$, $\forall i \in M$ $K_i = K_M$, Q монотонно.

Существует ли KO_M такая, что $\forall q \in Q$ $WG_M(k) = k_M$ при условии, что $\forall S \in G_M$ $K_S = K_M$ и $W_{S,L}(\cdot)$ монотонна?

Множество всех КО, являющихся решением задачи 1 для заданного обучающего набора Q обозначим $\mathrm{KO}_M(Q)$. Задача 1 может иметь много решений или не иметь решения даже в очень простых ситуациях.

В случае, когда задача 1 не может быть решена для какого-либо обучающего набора, то возможны следующие ее модификации.

Во-первых, можно говорить о задаче частичной реализации (аппроксимации) обучающего набора путем поиска КО, максимизирующей некий критерий, в частности критерий относительного результата $U_Q(KO_M) = \#\tilde{Q} / \#Q$, где $\tilde{Q} = \{q \in Q : WG_M(k) = k_M\}$. Для KO_M, реализующей Q, $U_Q(KO_M) = 1$, аппроксимирующей – $U_Q(KO_M) < 1$.

Задача 2. (аппроксимация обучающего набора КО с единой шкалой). Пусть $m > 2$, $K_M \in N$, $\forall i \in M$ $K_i = K_M$, Q монотонно. Необходимо найти KO_M такую, что $U_Q(KO_M) \to \max$ при условии, что $W_{S,L}(\cdot)$ монотонна?

Во-вторых, можно говорить о реализации обучающего набора с помощью КО, в которой шкалы на промежуточных уровнях не зафиксированы. При этом, традиции работы с механизмом комплексного оценивания рекомендуют сохранять свойство монотонности матричных сверток [2], что приводит к следующей постановки задачи реализации:

Задача 3. (Идентификация КО с монотонными матричными свертками). Пусть $m > 2$, $K_M \in N$, $\forall i \in M$ $K_i = K_M$, Q монотонно.

Существует ли KO_M такая, что $\forall q \in Q$ $WG_M(k) = k_M$ при условии, что $W_{S,L}(\cdot)$ монотонна?

Очевидно, что при этом можно ставить задачу минимального расширения множеств возможных значений оценок для промежуточных узлов дерева.

3. Алгоритм проверки реализуемости

В рамках доклада описывается алгоритм решения задачи 1. Суть предлагаемого в работе алгоритма состоит в последовательном разбиение множества критериев на два подмножества с проверкой, можно ли при таком разбиении в рамках заданного «алфавита» описать все элементы обучающего набора монотонной матрицей свёртки. Ключевой конструкцией в алгоритме является процедура разбиения множества Q на непересекающиеся множества с двухмерной адресацией на основе подмножества критериев $S \subset M$, синтезируя $W_{M,S}(\cdot)$. Кроме того, процедура позволяет определить «минимальный алфавит», описывающий элементы множества Q на основе интегральных оценок по подмножествам критериев S и $M \setminus S$. Если для некоторого $S \subset M$ $\#K_S \le \#K_M$ и $\#K_{M \setminus S} \le \#K_M$, то показывается, что данное разбиение допустимо для решения задачи 1 и $\{S, M \setminus S\} \subseteq G_M$, $W_{M,S}(\cdot) \in WG_M(\cdot)$. Далее процедура добавляется итеративно для всех подмножеств критериев, добавляемых в G_M до тех пор, пока не дойдет до листьев дерева – $S \in G_M : \#S = 1$. На каждой итерации для конкретной $S \in G_M : \#S > 1$ с помощью алгоритма проверяется, существует ли такое $L \subset S$, что по результатам процедуры

$\# K_L \leq \# K_M$ и $\# K_{S \setminus L} \leq \# K_M$ и синтезируется соответствующее $W_{S,L}(\cdot)$. Если ответ положительный, то показывается, что данное разбиение допустимо для решения задачи 1 и $G_M = G_M \cup \{L, S \setminus L\}$, и $WG_M(\cdot) = WG_M(\cdot) \cup W_{S,L}(\cdot)$.

4. Заключение

В докладе приведен алгоритм синтеза механизма комплексного оценивания на основе обучающего набора данных, решающий задачу 1. Обсуждается гипотеза о единственности получаемого с помощью алгоритма КО $KO_M(\cdot) = (G_M, WG_M(\cdot))$ для случая, если обучающий набор является полным - в Q представлены векторами со всеми возможными оценками по всем критериям. Обсуждается применимость алгоритма для решения задач 2 и 3 в случае, если задача 1 решения не имеет и применимость алгоритма для решения актуальных прикладных задач [5,6].

Работа выполнена при частичной поддержке РФФИ, проект 18-07-01258.

Список литературы

1. Трапезников В.А., Гореликов Н.И., Бурков В.Н., Зимоха В.А., Толстых А.В., Черкашин А.М., Цыганов В.В. Комплексный подход к управлению научно-техническим прогрессом в отрасли // Вестник АН СССР. 1983. № 3. С. 33-43.
2. Бурков В.Н., Новиков Д.А. Как управлять проектами. М.: Синтег, 1997. 188 с.
3. Казакова Е.А., Курочка П.Н. Автоматизированное построение матричных процедур комплексного оценивания на основе оптимизационного подхода // Вестник Воронежского государственного технического университета. 2010. Т. 6. С. 145–149.
4. Rokach L., Maimon O. Top-down induction of decision trees classifiers-a survey // IEEE Transactions on Systems, Man, and Cybernetics. Part C (Applications and Reviews). 2005. Vol. 35, No. 4. P. 476-487.
5. Пуликовский К.Б., Щепкин А.В. Комплексная оценка соответствия опасных производственных объектов требованиям безопасности // Безопасность труда в промышленности. 2007. № 4. С. 2-7.
6. Korgin N.A., Rozhdestvenskaya S.M. Concordant Approach for R&D Projects' Evaluation and Ranking // 2017 IEEE 11th International Conference on Application of Information and Communication Technologies (AICT). 2017. Vol. 2. P. 358-362.