Коллективные решения, ансамбли

Лектор – ***Сенько Олег Валентинович***

Курс «Математические методы обучения»

Содержание лекции

Задачи прогнозирования некоторой величины Y с помощью предикторов $Z_1, \dots, Z_M$. Под предикторами понимаются некоторые алгоритмы, прогнозирующие Y. Наборы алгоритмов, над которыми строятся коллективные решения в литературе по машинному обучению принято называть ансамблями. Пусть

$$Z_{ccp} = \sum_{i=1}^{L} c_i Z_i,$$

где

$$\sum_{i=1}^{L} c_i = 1,$$

$$c_i \geq 0, i = 1, \dots, L$$

Введём обозначение:
$\delta_i = \mathbb{E}(Y - Z_i)^2 -$ ошибка предиктора Z_i.

$$\sum_{i=1}^{L} \delta_i =$$

$$= \mathbb{E}\{\sum_{i=1}^{L} c_i(Y - Z_{ccp} + Z_{ccp} - Z_i)^2\} =$$

$$= \mathbb{E}\{(Y - Z_{ccp})^2 + \sum_{i=1}^{L} c_i(Z_{ccp} - Z_i)^2 + 2(Y - Z_{ccp})\sum_{i=1}^{L}(Z_i - Z_{ccp})\}$$

Структура ошибки Z_{ccp}

Однако

$$\sum_{i=1}^{L}(Z_i - Z_{ccp}) = 0$$

по определению Z_{ccp}
Следовательно

$$\delta Z_{ccp} = \mathbb{E}(Y - Z_{ccp})^2 = \sum_{i=1}^{L}\delta_i -$$

$$-\mathbb{E}[\sum_{i=1}^{L} c_i(Z_{ccp} - Z_i)^2]$$

Структура ошибки Z_{ccp}

Рассмотрим слагаемое

$$-\sum_{i=1}^{L} c_i (Z_{ccp} - Z_i)^2$$

Ннетрудно показать, что оно может быть представлено в виде

$$Z_{ccp}^2 - \sum_{i=1}^{L} c_i Z_i^2 = \sum_{i'=1}^{L} \sum_{i''=1}^{L} c_{i'} c_{i''} Z_{i'} Z_{i''} - \sum_{i=1}^{L} c_i Z_i^2$$

Воспользуемся равенством

$$Z_{i'}Z_{i''} = -\frac{1}{2}\{(Z_{i'} - Z_{i''})^2 -$$

$$-Z_{i'}^2 - Z_{i''}^2\} =$$

Откуда следует

$$\sum_{i'=1}^{L}\sum_{i''=1}^{L} c_{i'}c_{i''}Z_{i'}Z_{i''} - \sum_{i=1}^{L} c_i Z_i^2 =$$

$$= -\frac{1}{2}\sum_{i'=1}^{L}\sum_{i''=1}^{L} c_{i'}c_{i''}(Z_{i'} - Z_{i''})^2 +$$

$$+\frac{1}{2}\sum_{i'=1}^{L} c_{i'}Z_{i'}^2 + \frac{1}{2}\sum_{i''=1}^{L} c_{i''}Z_{i''}^2 - \sum_{i=1}^{L} c_i Z_i^2$$

Откуда следует

$$-\sum_{i=1}^{L} c_i (Z_{ccp} - Z_i)^2 = -\frac{1}{2} \sum_{i'=1}^{L} \sum_{i''=1}^{L} c_{i'} c_{i''} (Z_{i'} - Z_{i''})^2$$

Введём обозначение

$$\rho_{i'i''} = \mathbb{E}(Z_{i'} - Z_{i''})^2$$

Структура ошибки Z_{ccp}

$$\delta Z_{ccp} = \sum_{i=1}^{L} \delta_i - \frac{1}{2} \sum_{i'=1}^{L} \sum_{i''=1}^{L} c_{i'} c_{i''} \rho_{i' i''}$$

Лекция 9

Структура ошибки выпуклых комбинаций, комитетные методы, логическая коррекция

Лектор – *Сенько Олег Валентинович*

Курс «Математические основы теории прогнозирования»
4-й курс, III поток

Содержание лекции

1. Впуклые комбинации алгоритмов
2. Комитетные методы
3. Наивный байесовский корректор
4. Логическая коррекция
5. Алгебраическая коррекция

Ансамбли алгоритмов

Использование различных методов прогнозирования (распознавания), а также различных обучающих выборок или подмножеств признаков позволяет получить набор прогнозирующих (распознающих) алгоритмов: $A_1, \ldots, A_r$ Можно попытаться увеличить обобщающую способность за счёт выбора алгоритма с минимальной оценкой ошибки прогнозирования. Однако нередко более эффективной процедурой является вычисление прогноза с использованием всех алгоритмов из $A_1, \ldots, A_r$. Использование коллектива (ансамбля) алгоритмов, которые строятся с помощью различных методов позволяет использовать при прогнозировании различные принципы экстраполяции, лежащих в основе этих методов. Статистическое обоснование использованию ансамбля алгоритмов даёт анализ ошибки выпуклой комбинации прогнозов, вычисляемых членами ансамбля. Предположим, что алгоритмы ансамбля $A_1, \ldots, A_r$ вычисляют прогноз переменной Y.

Выпуклые комбинации

Пусть f_i - прогноз, вычисляемый алгоритмом A_i. Тогда

$$\Delta_i = E_\Omega(Y - f_i)^2$$

вляется математическим ожиданием квадрата ошибки прогнозировани для A_i , . Введём обозначение $\rho_{i'i''}$ для математического ожидания квадрата отклонения друг от друга прогнозов, вычисляемых алгоритмами $A_{i'}$ и $A_{i''}$. То есть

$$\rho_{i'i''} = E_\Omega(f_{i'} - f_{i''})^2.$$

Пусть $c_1, \ldots, c_r$ -положительные коэффициенты такие, что $\sum_{i=1}^{r} c_i = 1$ Обозначим через $\hat{f}$ выпуклую комбинацию прогнозов, вычисляемых алгоритмами ансамбля $A_1, \ldots, A_r$. То есть

$$\hat{f} = \sum_{i=1}^{r} c_i f_i.$$

Для ошибки выпуклой комбинации справедливо выражение

$$\hat{\Delta} = E_{\Omega}(Y - \hat{f})^2 = \sum_{i=1}^{r} c_i \Delta_i - \frac{1}{2} \sum_{i'=1}^{r} \sum_{i''=1}^{r} c_{i'} c_{i''} \rho_{i'i''} \tag{1}$$

Принимая во внимание, что все квадратичные отклонения $\rho_{i'i''}$ всегда неотрицательны, а коэффициенты $c_1, \ldots, c_r$ положительны получаем неравенство

$$\hat{\Delta} \leq \sum_{i=1}^{r} c_i \Delta_i.$$

Иными словами математическое ожидание квадрата ошибки выпуклой комбинации всегда не превышает аналогичную выпуклую комбинацию математических ожиданий квадратов ошибок отдельных алгоритмов ансамбля.

Рассмотрим, случай, когда все алгоритмы участвуют в построении коллективного решения равноправно. В этом случае $c_i = \frac{1}{r}$, $i = 1, \dots, r$. Выпуклая комбинация становится просто средним значением

$$\hat{f} = \frac{1}{m} \sum_{i=1}^{r} f_i.$$

Математическое ожидание квадрата ошибки усреднённого по ансамблю прогноза вычисляется по формуле

$$\hat{\Delta} = E_\Omega (Y - \hat{f})^2 = \frac{1}{m} \sum_{i=1}^{r} \Delta_i - \frac{1}{2} \frac{1}{m^2} \sum_{i'=1}^{r} \sum_{i''=1}^{r} c_{i'} c_{i''} \rho_{i'i''} \qquad (2)$$

Таким образом математическое ожидание квадрата ошибки усреднённого по ансамблю прогноза представляет собой разницу между средней по ансамблю величиной математического ожидания квадрата ошибки и средней величиной квадратичного отклонения между прогнозами вычисляемыми различными алгоритмами.

Рассмотрим сначала несколько простейших эвристических методов принятия коллективных решений. Предположим, что у нас есть ансамбль алгоритмов распознавания $A_1, \dots, A_r$, которые были использованы для классификации некоторого объекта s^*.

Голосование по большинству. Простейшим комитетным методом является является метод голосования по большинству, относящий объект к тому классу, к которому он был присвоен относительным большинством алгоритмов.

Использование вещественных оценок за классы. Напомним, что произвольный рспознающий алгоритм является комбинацией распознающего оператора, вычисляющего оценки за классы и решающего правила, производящего классификацию по оценкам, вычисленным распознающим оператором. Предположим, что $\Gamma_l^i(*)$ - оценка за класс , вычисляемая алгоритмом A_i . Коллективное решение может строится путём вычисления коллективных оценок за классы через оценки $\Gamma_l^i(*)$, соответствующие отдельным алгоритмам. При этом могут использоваться различные варианты вычисления

1) Коллективная оценка за класс K_l вычисляется как среднеарифметическое оценок, вычисляемых алгоритмами из ансамбля $\{A_1, \ldots, A_r\}$:

$$\Gamma_l^{av}(s^*) = \sum_{j=1}^{r} \Gamma_l^j(s^*).$$

2) Коллективная оценка за класс K_l вычисляется как вычисляется как минимум всех оценок за данный класс полученных алгоритмами из ансамбля $\{A_1, \ldots, A_r\}$:

$$\Gamma_l^{min}(s^*) = \min_{j \in \{1,\ldots,r\}} \Gamma_l^j(s^*).$$

3) Коллективная оценка за класс K_l вычисляется как вычисляется как максимум всех оценок за данный класс полученных алгоритмами из ансамбля $\{A_1, \ldots, A_r\}$:

$$\Gamma_l^{min}(s^*) = \max_{j \in \{1,\ldots,r\}} \Gamma_l^j(s^*).$$

4) Еще одним употребительным способом построения комитетного решения является использование произведений оценок, вычисляемых алгоритмами из ансамбля $\{A_1, \ldots, A_r\}$:

$$\Gamma_l^{av}(s^*) = \prod_{j=1}^{r} \Gamma_l^j(s^*).$$

К достоинствам комитетных методов относится их простота, высокая быстродействие. Для применения этих методов не требуется никакой дополнительной процедуры обучения, что позволяет сразу переходить к распознаванию объектов комитетом обученных алгоритмов.

Подобными же достоинствами обладает другой известный метод построения коллективных решений – «Наивный байесовский классификатор», который является статистическим методом, основанном на оценках вероятностей принадлежности объекта классам в зависимости от результатов классификации отдельными алгоритмами. Предположим, что алгоритмы $\{A_1, \ldots, A_r\}$ отнесли объект s^*в классы $K_{J(1)}, \ldots, K_{J(r)}$ соответственно. Факт отнесения объекта s в класс K_i алгоритмом A_j далее будем обозначать $A_j(s) = \mathrm{Pt}_i(s)$, где $\mathrm{Pt}_i(s)$ является предикатом, обозначающим отнесение s в класс K_i. Наибольшую точность распознавания обеспечивает байесовский классификатор, относящий объект в класс K_{i^*}, для которого максимальной является условная вероятность

$$P[s^* \in K_{i^*} \mid A_1(s^*) = \mathrm{Pt}_{J(1)}(s^*), \ldots, A_r(s^*) = \mathrm{Pt}_{J(r)}(s^*)] \qquad (3)$$

Условная вероятность (1) для класса K_i может быть вычислена по формуле Байеса

$$P[s^* \in K_i \mid A_1(s^*) = \mathrm{Pt}_{J(1)}(s^*), \ldots, A_r(s^*) = \mathrm{Pt}_{J(r)}(s^*)] =$$

$$= \frac{P(K_i)P[A_1(s^*) = \mathrm{Pt}_{J(1)}(s^*), \ldots, A_r(s^*) = \mathrm{Pt}_{J(r)}(s^*) \mid s^* \in K_i]}{P[A_1(s^*) = \mathrm{Pt}_{J(1)}(s^*), \ldots, A_r(s^*) = \mathrm{Pt}_{J(1)}(s^*)]}.$$

Условная вероятность

$$P[A_1(s^*) = \mathrm{Pt}_{J(1)}(s^*), \ldots, A_r(s^*) = \mathrm{Pt}_{J(1)}(s^*) \mid s^* \in K_i]$$

может быть оценена исходя из гипотезы о независимости входящих в ансамбль классификаторов. То есть

$$P[A_1(s^*) = \mathrm{Pt}_{J(1)}(s^*), \ldots, A_r(s^*) = \mathrm{Pt}_{J(r)}(s^*) \mid s^* \in K_i] =$$

$$= \prod_{j=1}^{r} P[A_1(s^*) = \mathrm{Pt}_{J(j)}(s^*) \mid s^* \in K_i].$$

В качестве оценок вероятностей

$$P(K_1), \dots, P(K_l)$$

$$P[A_1(s^*) = \mathrm{Pt}_{J(j)}(s^*) \mid s^* \in K_i],$$

при

$$j = 1, \dots, r, i = 1, \dots, r$$

могут быть использованы соответствующие доли объектов обучающей выборки. Отметим, что вероятность

$$P[A_1(s^*) = \mathrm{Pt}_{J(1)}(s^*), \dots, A_r(s^*) = \mathrm{Pt}_{J(r)}(s^*)]$$

является одинаковым для всех классов множителей, который может не учитываться при вычислении окончательного решения.

Комитетные методы и наивный байесовский классификатор являются простейшими методами коллективной коррекции, не учитывающих взаимодействие алгоритмов ансамбля или их относительную эффективность. Требование повышения обобщающей способности ансамбля за счёт более полного учёта его структуры и использования возможностей лежащих в его основе эвристик. привело к созданию средств алгебраической и логической коррекции. Методы логической коррекции учитывают только окончательные результаты классификации. Пусть у нас имеется некоторая выборка $\widetilde{S}_c = \{s_1, \dots, s_q\}$ объектов, принадлежащих классам $K_1, \dots, K_L$, по которой мы собираемся произвести коррекцию. Данной выборке может быть сопоставлена информационная матрица $\|\alpha_{li}\|_{L\times q}$, , где $\alpha_{li} = 1$ при $s_i \in K_l$ и $\alpha_{ij} = 0$ в противном случае.

.

Выборке $\widetilde{S}_c$ может быть также сопоставлен набор матриц

$$\{\|\beta_{li}^{j}\|_{L\times q} \mid j = 1, \ldots, r\}.$$

Элемент $\beta_{li}^{j} = 1$, если $A_j(s_i) = \mathrm{Pt}_l(s_i)$, и $\beta_{li}^{j} = 0$ в противном случае. Поиск оптимального логического корректора сводится к поиску для каждого класса такой логической функции $F_l(z_1, \ldots, z_r)$ от булевых переменных $z_1, \ldots, z_r$, чтобы равенство

$$F_l[\beta_{li}^{g(1)}, \ldots, \beta_{li}^{g(r)}] = \alpha_{li}$$

выполнялось для возможно большего числа объектов выборки $\widetilde{S}_c$. Функция $g(i)$ устанавливает связь между переменными $z_1, \ldots, z_r$ и алгоритмами $A_1, \ldots, A_r$. Использование g позволяет учитывать информативность алгоритмов для оценки принадлежности распознаваемых объектов классу K_l, через их место в логической функции.

Предположим, что отсутствуют противоречия типа существования в выборке $\widetilde{S}_c$ объектов $s_{i'}$ и $s_{i''}$ с неодинаковыми $\alpha_{li'}$ и $\alpha_{li''}$, которые однако одинаково классифицируются алгоритмами ансамбля. То есть $\beta^j_{li'} = \beta^j_{li''}$ при $j = 1, \dots, r$.
В случае, если задана какая-либо функция g, а множество векторов

$$\{(\beta^1_{li}, \dots, \beta^r_{li}) \mid i = 1, \dots, r\}$$

включает всё множество вершин единичного куба $\mathbb{E}^r$, логическая функция F_l оказывается полностью определённой. В противном случае задача построения логического корректора включает в себя задачу доопределения логической функции естественным путём заданной на выборке $\widetilde{S}_c$ на весь единичный куб $\mathbb{E}^r$.

Одним из способов логической корреции является построение монотонных корректоров, которое сводится к поиску такой функции g, что логическая функция F_l, правильно вычислющая элнменты информационной матрицы, является монотонной. То есть ищется функция $g(i)$, которая
а) удовлетворяет равенству

$$F_l[\beta_{li}^{g(1)}, \dots, \beta_{li}^{g(r)}] = \alpha_{li}$$

при $i = 1, \dots, q$;
б) для любых векторов $(z'_1, \dots, z'_r)$ и $(z''_1, \dots, z''_r)$, удовлетворяющих условию

$$(z'_1, \dots, z'_r) \succeq (z''_1, \dots, z''_r)$$

выполняется неравенство

$$F_l(z'_1, \dots, z'_r) \succeq F_l(z''_1, \dots, z''_r)$$

.

Построение монотонных корректоров сводится к следующей схеме. В исходном наборе $A_1, \dots, A_r$ для каждого класса K_l выбирается поднабор $A_{f(1)}, \dots, A_{f(k)}$. Объект s относится монотонным логическим корректором в класс K_l в том и только в том случае, если он отнесён в K_l всеми алгоритмами из $A_{f(1)}, \dots, A_{f(k)}$ и ещё одним алгоритмом из набора $A_1, \dots, A_r$, который не принадлежит $A_{f(1)}, \dots, A_{f(k)}$.

Алгебраическая коррекция

Универсальным способом построения оптимального распознающего алгоритма по набору исходных алгоритмов $A_1, \ldots, A_r$ является использование алгебраических методов коррекции. В отличие от логических методов коррекции алгебраические методы используют не только окончательные результаты классификации, содержащиеся в матрицах $\|\beta_{li}^j\|_{L\times q}$, но также матрицы оценок $\|\gamma_{li}^j\|_{L\times q}$, вычисляемые операторами $R_1, \ldots, R_k$, входящими в алгоритмы $A_1, \ldots, A_r$. Элемент γ_{li}^j является оценкой объекта s_i за класс K_l , вычисляемая оператором R_j, $i = 1, \ldots, q$,$l = 1, \ldots, L$, $j = 1, \ldots, r$. Основы теории алгебраической коррекции были разработаны Ю.И.Журавлёвым в 1976-1978 годах. Задача распознавания в алгебраической теории рассматривается как задача построения по начальной информации I о классах $K_1, \ldots, K_L$ для предъявленной для распознавания выборки $\widetilde{S}_c = \{s_1, \ldots, s_q\}$ правильной информационной матрицы $\|\alpha_{li}^j\|_{L\times q}$. .

Последнюю задачу мы будем называть задачей $\mathrm{Z}(I, \widetilde{S}_c, \mathrm{Pt}_1, \dots, \mathrm{Pt}_L)$ или просто задачей Z. Примером начальной информации о классах является таблица признаковых описаний эталонных объектов классов и их информационная матрица. Предположим, что у нас имеется множество алгоритмов $\{A\}$, переводящих пару $(I, \widetilde{S}_c)$ в матрицы $\|\beta_{li}^j\|_{L\times q}$, составленные из элементов $\{0, 1, \Delta\}$, где значения 1 и 0 как и раньше соотвествуют истинности или ложности предикатов, вычисленными алгоритмами из множества $\{A\}$, значение Δ соответствует отказу от вычисления значения предиката.

Определение 1. Алгоритм A называется корректным для задачи Z, если выполнено равенство

$$A(I, \widetilde{S}_c, \mathrm{Pt}_1, \dots, \mathrm{Pt}_L) = \|\alpha_{li}\|_{L\times q}.$$

Алгоритм, не являющийся корректным для задачи Z , называется некорректным.

Совокупность $\{A\}$ состоит из вообще говоря некорректных алгоритмов. Алгебраический подход к решению задач распознавания включает в себя введение алгебраических операций над алгоритмами из $\{A\}$, позволяющих строить корректные алгоритмы по наборам алгоритмов из $\{A\}$. Поскольку каждый распознающий алгоритм может быть представлен как последовательное выполнение распознающего оператора и решающего правила, множеству $\{A\}$ соответствуют множества операторов $\{R\}$ и множество решающих правил $\{C\}$. Каждый из операторов из множества $\{R\}$ вычисляет для задачи Z матрицу оценок за классы

$$R^*(I, \widetilde{S}_c) = \|\gamma^*_{li}\|_{L\times q}$$

На множестве операторов $\{R\}$ вводятся операции сложения, умножения и умножения на скаляр.

Предположим, что R' и R'' являются операторами из $\{R\}$. При этом $R'(I,\widetilde{S}_c) = \|\gamma'_{li}\|_{L\times q}$ и $R''(I,\widetilde{S}_c) = \|\gamma''_{li}\|_{L\times q}$. Пусть b является некоторой скалярной величиной. Операция умножения на скаляр преобразует оператор R' в оператор $(b \bullet R')$, задаваемый формулой

$$(b \bullet R')(I,\widetilde{S}_c) = \|b\gamma'_{li}\|_{L\times q}, \tag{4}$$

Сумма операторов $(R' + R'')$ задаётся формулой

$$(R' + R'')(I,\widetilde{S}_c) = \|\gamma'_{li} + \gamma''_{li}\|_{L\times q}, \tag{5}$$

Произведение операторов $(R' \bullet R'')$ задаётся формулой

$$(R' + R'')(I,\widetilde{S}_c) = \|\gamma'_{li} \cdot \gamma''_{li}\|_{L\times q}, \tag{6}$$

Использование операций (4)-(6) позволяет строить новые распознающие операторы, являющиеся полиномами от операторов из исходного множества вида

$$\sum_{i=1}^{N_p} a_i [R_{t(1,i)} \bullet \ldots \bullet R_{t(k(i),i)}]$$

Функция $t(j,i)$ указывает на оператор, находящийся в позиции j слагаемом с номером i, k_i -число сомножителей в слагаемом с номером i. Очевидно, что замыкание $\mathbf{L}\{R\}$ множества операторов $\{R\}$ относительно операций (4) и (5) является линейным векторным пространством. Обозначим через $\mathbf{U}\{R\}$ алгебраическое замыкание множества $\{R\}$ относительно операций (4)-(6).
Рассмотрим условия, существования корректного алгоритма для некоторой задачи $\mathrm{Z}(I, \widetilde{S}_c, \mathrm{Pt}_1, \ldots, \mathrm{Pt}_L)$.

. **Определение 2.** Если множество матриц $\{R(I,\widetilde{S}_c)\}$, где операторы пробегают множество $\{R\}$, содержит базис в пространстве числовых матриц размерности $L \times q$, то задача $\mathrm{Z}(I,\widetilde{S}_c,\mathrm{Pt}_1,\ldots,\mathrm{Pt}_L)$ называется полной относительно $\{R\}$.

Определение 3. Решающее правило C называется корректным , если для всякой выборки длины q существует хотя бы одна числовая матрица $\|\gamma'_{li}\|_{L\times q}$ такая, что

$$C(\|\gamma'_{li}\|_{L\times q}) = \|\alpha_{li}\|_{L\times q}$$

Пусть $\{A\}$ является множество алгоритмов вида $A = R \otimes C^*$, где $R \in \{R\}$, C^* - некоторое корректное решающее правило.

Определение 4. Множества алгоритмов вида $A = R \otimes C^*$, будут обозначаться $\mathbf{L}\{A\}$ и $\mathbf{U}\{A\}$, если $R \in \mathbf{L}\{R\}$ и $R \in \mathbf{U}\{R\}$ соответственно.

Пусть C^* - некоторое корректное решающее правило.
Теорема 1. Если множество $\{Z\}$ состоит лишь из задач , полных относительно $\{R\}$, то линейное замыкание $\mathbf{L}\{A = R \otimes C^*\}$, является корректным относительно $\{Z\}$.
Доказательство. Пусть M является матрицей, которая может быть переведена решающим правилом C^* в информационную матрицу $\|\alpha_{li}\|_{L\times q}$. Существование матрицы M следует из корректности решающего правила C^*. При фиксированном q базис в пространстве числовых матриц размерности $L \times q$ состоит из Lq матриц $M_1, \ldots, M_{Lq}$.Тогда существуют такие числа $c_1, \ldots, c_{Lq}$, что $M = \sum_{i=1}^{Lq} c_i M_i$.

В том случае, если матрицы $M_1, \ldots, M_{Lq}$ построены из $\{I, \widetilde{S}_c\}$ с помощью операторов $R_1, \ldots, R_{Lq}$ из $\{R\}$, корректный алгоритм A_{corr} можетбыть представлен в виде

$$A_{corr} = \sum_{i=1}^{Lq} c_i R_i \otimes C^*$$

Теорема доказана.
Следствие 1. Пусть $\{A\}$ - совокупность некорректных алгоритмов, $\{R\}$ - соответствующее множество операторов, C^* - некоторое фиксированное корректное решающее правило. Тогда $\mathbf{L}\{A = R \otimes C^*\}$ является корректным относительно $\{Z\}$, если $\{Z\}$ состоит из задач, полных относительно $\{R\}$.

Следствие 2. Пусть выполнены все условия следствия 1. Тогда $\mathbf{U}\{A\}$ является корректным относительно $\{\mathrm{Z}\}$, если $\{\mathrm{Z}\}$ состоит из задач, полных относительно замыкания $\mathbf{U}\{A\}$.

Линейные и алгебраические замыкания могут строится не только над конечными наборами заранее обученных алгоритмов, но также и над множествами алгоритмов, принадлежащих некоторой модели и имеющих в общем случае мощность континуума. Рассмотрим в качестве примера рассмотрим один из вариантов модели алгоритмов вычисления оценок, в котором оценки за класс K_l вычисляются по формуле

$$\Gamma_l(s^*) = \sum_{s_\mu \in K_l} \sum_{\omega \in \Omega} (\sum_{j=1}^{n} p_j \omega_j) B_\omega(s^*, s_\mu, \varepsilon) + \tag{7}$$

$$+ \sum_{s_\mu \in \widetilde{S} \setminus K_l} \sum_{\omega \in \Omega} (\sum_{j=1}^{n} p_j \omega_j) \bar{B}_\omega(s^*, s_\mu, \varepsilon)$$

В формуле (7) используются следующие обозначения

- $\bar{B}_{\boldsymbol{\omega}}(s^*, s_\mu, \boldsymbol{\varepsilon}) = 1 - B_{\boldsymbol{\omega}}(s^*, s_\mu, \boldsymbol{\varepsilon})$ является функцией антиблизости объекта к эталону по опорному множеству, описываемому бинарным характеристическим вектором $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)$;
- $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)$- вектор положительных пороговых коэффициентов, задающих близость объектов по каждому из признаков;
- $(p_1, \ldots, p_n)$-вектор положительных параметров, характеризующих важность признаков,
- $(\gamma_1, \ldots, \gamma_n)$ вектор положительных параметров, характеризующих важность объектов.

Для того, чтобы описать условия существования корректного алгоритма в алгебраическом замыкании подмножества алгоритмов вычисления оценок введём дополнительные определения. Пусть $\widetilde{M}$ - некоторое множество допустимых объектов.

Определение 5. Объекты s_u и s_ν с описаниями $(x_1^u, \ldots, x_n^u)$ и $(x_1^\nu, \ldots, x_n^\nu)$ называются изоморфными относительно множества $\widetilde{M}$, если для произвольного объекта s с описанием $(a_1, \ldots, a_n) \in \widetilde{M}$ выполняются равенство

$$(| x_1^u - a_1 |, \ldots, | x_n^u - a_n |) = (| x_1^\nu - a_1 |, \ldots, | x_n^\nu - a_n |)$$

Нетрудно видеть что корректный алгоритм в рамках модели вычисления оценок по формуле (7) не может существовать для задачи $\mathrm{Z}(I, \widetilde{S}_c, \mathrm{Pt}_1, \ldots, \mathrm{Pt}_L)$ случаях, когда оказываются выполненными следующие два условия.

- в выборке $\widetilde{S}_c$ существует два объекта s' и s'', изоморфных относительно выборки эталонов $\widetilde{S}_e$, которая вместе со своей информационной матрицей образует начальную информацию I;
- объекты s' и s'' принадлежат двум непересекающимся классам. Действительно в этом случае векторы оценок $[\Gamma_1(s'), \ldots, \Gamma_L(s')]$ и $[\Gamma_1(s''), \ldots, \Gamma_L(s'')]$, вычисляемые произвольным оператором из модели АВО будут одинаковы. Следовательно никакое множество операторов модели АВО не может вычислять базис в пространстве вещественных матриц размера $L \times q$.

Будем называть задачу $\mathrm{Z}(I, \widetilde{S}_c, \mathrm{Pt}_1, \ldots, \mathrm{Pt}_L)$ регулярной при выполнении трёх условий:

- никакие два класса полностью не совпадают, т.е. $K_{l'} \neq K_{l''}$ при $l' \neq l''$;
- никакие два объекта из $\widetilde{S}_c$ не являются изоморфными относительно выборки эталонов $\widetilde{S}_e$;
- $\widetilde{S}_e \bigcap \widetilde{S}_c = \emptyset$

Справедлива следующая теорема.
Теорема 2. Алгебраическое замыкание подкласса алгоритмов модели АВО, в которой оценки за классы вычисляются по формуле (5) корректно над множеством регулярных задач.