УДК 519.95

СИНТЕЗ СОГЛАСОВАННЫХ ЛИНЕЙНЫХ ОПТИМИЗАЦИОННЫХ МОДЕЛЕЙ ПО ПРЕЦЕДЕНТНОЙ ИНФОРМАЦИИ: ПОДХОД НА ОСНОВЕ КОЛМОГОРОВСКОЙ СЛОЖНОСТИ

© 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@tnu.crimea.ua*

Abstract. In the paper, approach is expounded to the analysis of the linear optimization models built on precedent initial learning information. In supposition, that all numerical parameters are rational and limited by the bit net of computer, estimations of Kolmogorov's complexity and nonrandomness of extraction of model as empirical regularity are got from the sample.

Введение

Линейные оптимизационные модели во многих случаях хорошо описывают процессы экономического планирования производства, распределения ресурсов, размещения объектов. Активно развивающееся в математике направление, связанное с алгоритмической сводимостью, позволяет увидеть, как часто возникающие на практике задачи сводятся к задачам линейной оптимизации. В то же время на этапе внедрения линейных моделей в производственное управление выявились две основные проблемы, связанные, главным образом, с ростом размерности. Первая — сложность сбора данных для фиксации модели, вторая — неадекватность модели реальному процессу или объекту вследствие неполноты и/или неадекватности данных.

Основная идея, на которой основан развиваемый в статье подход, состоит в том, что *линейная модель должна быть согласована с реальной, фактической информацией об объекте моделирования. Такой фактической информацией являются прецеденты — наблюдения, зафиксированные в процессе функционирования объекта.*

Многие задачи линейного программирования с вещественными неотрицательными переменными имеют вид

$$\begin{cases} \max f(X) \\ AX^T \leqslant B \end{cases}, \tag{1}$$

где $X = (x_1, ..., x_n)$, $A = \|a_{i,j}\|_{m\times n}$, $B = (b_1, ..., b_m)^T$, $x_i \geqslant 0$, $a_{i,j} \in \mathbb{R}$, $b_i \in \mathbb{R}$, $f : \mathbb{R}^n \to \mathbb{R}$. Будем называть $\Omega = \{X : AX^T \leqslant B\}$ множеством допустимых решений задачи (1).

Задача, которая решается в данной статье, состоит в следующем. Предполагается, что для решения оптимизационной задачи вида (1) в некоторой проблемной области представлена только начальная информация в виде конечного набора точек из $\mathbb{R}^n$ вместе со значением неизвестной целевой функции в каждой точке (если это значение измеряно), а также логическим значением 1, если точка удовлетворяет всем требуемым ограничениям, которые определяются проблемной областью, или 0 — если эта точка ограничениям не удовлетворяет. Требуется: a) обосновать возможность применения модели линейного программирования (1) и б) — если применение линейной модели возможно, — синтезировать модель по предоставленной прецедентной начальной информации, т.е. построить матрицу $\hat{A}$ и вектор $\hat{B}$, определив при этом число линейных ограничений-неравенств $\hat{m}$, а также найти целевую функцию $\hat{f}$. Затем синтезированная модель

$$\begin{cases} \max \hat{f}(X) \\ \hat{A}X^T \leqslant \hat{B} \end{cases} \tag{2}$$

должна быть некоторым образом обоснована, т.е. требуется c) дать оценку адекватности синтезированной модели (2).

Формально начальная информация имеет вид набора троек $D = \{(X_k, y_k, \omega_k),\ k = 1, ..., l\}$. Будем называть эту информацию *обучающей выборкой или данными* D, а число l заданных точек — *длиной выборки.* $X_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}$ (y_k может быть не определено для части элементов выборки; для пропущенных значений будет использованться обозначение Δ); $\omega_k = 1$, если выборочная точка удовлетворяет ограничениям, и $\omega_k = 0$ — если она ограничениям не удовлетворяет. Обозначим $D_\oplus = \{X_k : \omega_k = 1\}$; $|D_\oplus| = l_\oplus$ — число допустимых решений, представленных в начальной информации; $D_\ominus = \{X_k : \omega_k = 0\}$; $|D_\ominus| = l_\ominus$ — число точек, не являющихся допустимыми решениями.

Множество допустимых решений модели (2) будет обозначаться $\hat{\Omega}_D = \{X : \hat{A}X^T \leqslant \hat{B}\}$.

Синтезированная модель (2) является аппроксимацией модели (1), которая полагается существующей и представленной неполной начальной информацией D.

Впервые постановка подобных задач была сделана В. Д. Мазуровым в работах [10] и [11]. Для их решения использовались итерационные методы, основанные на фейеровских отображениях [12].

Синтетический подход к решению линейных оптимизационных задач по частичной (прецедентной) информации был предложен и разрабатывался в работах [8, 6, 18, 15, 2] и был, в основном, основан на эвристиках. Исследования последних лет в области теории колмогоровской сложности и её применения в машинном обучении [1, 7, 19] определили интерес к дальнейшему исследованию проблем синтеза оптимизационных моделей по прецедентной информации.

1. Синтез линейной модели как машинное обучение

Применяя методы машинного обучения по прецедентной информации D, возможно использовать линейные итерационные модели, основанные на процедуре Розенблатта-Новикова, которые позволяют построить набор из $\hat{m}$ линейных гиперплоскостей, определяющих область допустимых решений $\hat{\Omega}_D$ в виде $\hat{L}_q(X) \leqslant \hat{b}_q$, $q = 1, ..., \hat{m}$, где $\hat{L}_q$ — линейные функции, а $\hat{b}_q$, $q = 1, ..., \hat{m}$, — пороговые вещественные значения. Коэффициенты этих функций и числа $\hat{b}_q$, $q = 1, ..., \hat{m}$, будут определять параметры $\hat{A}$ и $\hat{B}$ синтезированной модели (2). А число $\hat{m}$ полученных неравенств (вместе с заданным n) определит размерность матрицы $\hat{A}$ и вектора $\hat{B}$. Синтез целевой функции может быть осуществлён, например, по методу наименьших квадратов.

Такой подход был предложен в [8]. Интуитивно заложенный в алгоритм синтеза ограничений, описанного в этой работе, принцип «бритвы Оккама» (Occam's Razor), как будет показано ниже, может быть достаточно строго обоснован на базе понятия колмогоровской алгоритмической сложности.

Нужно подчеркнуть, что излагаемые в настоящей статье методы и алгоритмы рассчитаны на компьютерную реализацию. Это еще не раз будет подчёркиваться по ходу изложения материала.

2. Согласованная модель-гипотеза

Модель (1) с фиксированными значениями (параметрами) n, m, f, A, B является индивидуальным представителем класса $\mathscr{L}$ моделей линейного программирования. Синтезированная модель (2), как правило, не совпадает с истинной неизвестной моделью (1) и является её аппроксимацией.

Определение 1. Синтезированная модель (2) называется *согласованной (согласованной гипотезой)*, если при подстановке точек из обучающей выборки D в целевую функцию и линейные неравенства модели (2) выполняются условия: $\hat{f}(X_k) \equiv y_k$ и $[\hat{A}X_k \leqslant \hat{B}] \leftrightarrow \omega_k$ для всех $k = 1, .., l$, соответствующих полностью заданным (определенным) тройкам значений из D.

Согласованная линейная модель будет далее обозначаться $\mathscr{M}_D$.

Нужно заметить, что в широком семействе моделей $\mathscr{L}$ существует более узкий подкласс согласованных моделей $\mathscr{L}_D$. Выбор согласованной модели из $\mathscr{L}_D$ — сложная, некорректная по Адамару задача, требующая серьёзного обоснования.

Учитывая, что все излагаемые в статье методы и алгоритмы рассчитаны на компьютерную реализацию, *можно и нужно сузить рассматриваемые семейства согласованных моделей до конечных, ограничив представления чисел рациональными значениями, лежащими в отрезке, который определяется разрядной сеткой применяемого компьютера* .

3. Колмогоровская сложность модели линейной оптимизации

Комогоровская сложность может определяться как в терминах класса машин Тьюринга (MT), так и в терминах класса частично рекурсивных функций P_{pr}, поскольку эти два класса эквивалентны. Поэтому, если не возникает недоразумений, в дальнейшем один и тот же объект U может быть назван $MT\ U$ или функцией $U \in P_{pr}$. Тьюринговская интерпретация предпочтительнее, когда рассуждения проводятся в терминах строк и алфавитных отображений. Частично рекурсивная — когда необходимо использовать свойства алгоритмов как функций при доказательстве теорем.

Если $\mathscr{A} = \{\alpha_1, ..., \alpha_\mu\}$ — конечный алфавит, то $\mathscr{A}^*$ — множество любых строк любой конечной длины над алфавитом $\mathscr{A}$. Будем обозначать $p \in \{0,1\}^*$ — двоичные строки. Обозначаемые другими буквами строки могут состоять из символов иных, отличных от $\{0,1\}^*$ алфавитов. Длина произвольной строки s (число символов в ней) обозначается $|s|$. Символом Λ будет обозначаться пустая строка; $|\Lambda| = 0$.

(Префиксная) колмогоровская сложность произвольной строки s при заданной МТ U определяется как $K_U(s) = min\{\,|p| : U(p) = s\}$, где p — бинарная строка, которую назвают наименьшей длиной программы, вычисляющей s (или описания, кодирующего s). Поскольку универсальная МТ может моделировать любую другую МТ, выбор U в формуле $K_U(s)$ может изменить колмогоровскую сложность не более чем на константу, зависящую только от U [9]. Поэтому U в обозначении $K_U(s)$ часто опускается, и используется обозначение $K(s)$.

Нужно отметить, что существует МТ U_s такая, которая, получая вход в виде пустой строки, генерирует заданную строку s. Чтобы это понять, достаточно указать последовательность команд, определяющую U_s, в которой соответствующая последовательность символов вывода в правых частях идущих строго друг за другом команд, завершающихся командой остановки, в точности совпадает с s. Можно сказать, что в этом случае s «вмонтирована» в тело программы машины U_s. Но тогда получается, что в определении $K_U(s) = min\{|p| : U(p) = s\}$ машина U не может быть любой, поскольку для пустой входной строки Λ машина U_s выдаст s, и тогда $K_{U_s}(s) = 0$ для любой строки s. Будем называть МТ U_s *генератором* строки s, если эта машина для любой входной строки выдаёт строку s, и обозначать $\mathscr{U}_s$ — множество генераторов для всевозможных строк s.

Определение 2. *Точной колмогоровской сложностью произвольной строки* s называется

$$KC(s) = \min_{U \notin \mathscr{U}_s} (\min\{|p| : U(p) = s\}),$$

если равенство $U(p) = s$ выполняется для хотя бы одной МТ, не являющейся генератором, иначе полагается что $KC(s) = \infty$.

Введение определения точной колмогоровской сложности объясняется необходимостью избавиться от константы, связанной с переходом от одной оптимальной машины U к другой. Эта константа фигурирует в классическом определении и усложняет получение оценок неслучайности выбора моделей в задачах машинного обучения, основанных на принципе минимальной длины описания (MDL).

Определение 3. *Колмогоровской сложностью согласованной модели при заданной МТ U* называется

$$KC_U(\mathscr{M}_D) = \min\left\{|p| : U(p) = s_{\hat{C}} * s_{\hat{A}} * s_{\hat{B}}\right\},$$

где $s_{\hat{C}} * s_{\hat{A}} * s_{\hat{B}}$ — строка, являющаяся конкатенацией трёх строк, определяющих такие элементы модели $\mathscr{M}_D$, что

$$(\hat{f}_{\hat{C}}(X_k) \equiv y_k) \wedge ([\hat{A}X_k \leqslant \hat{B}] \leftrightarrow \omega_k), (X_k, y_k, \omega_k) \in D, \quad k = 1, ..., l,$$

а $\hat{f}_{\hat{C}}$ — целевая функция, определяемая набором коэффициентов $\hat{C}$.

Определение 4. *Колмогоровской сложностью согласованной системы линейных ограничений при заданной МТ U* называется

$$KC_U(\hat{\Omega}_D) = \min\left\{|p| : U(p) = s_{\hat{A}} * s_{\hat{B}}\right\},$$

где $s_{\hat{A}} * s_{\hat{B}}$ — строка, являющаяся конкатенацией двух строк, определяющих такие элементы модели $\hat{\Omega}_D$, что

$$([\hat{A}X_k \leqslant \hat{B}] \leftrightarrow \omega_k), (X_k, \Delta, \omega_k) \in D, \quad k = 1, ..., l.$$

Определение 5. *Точной колмогоровской сложностью семейства* $\mathfrak{M}_D$ согласованных моделей называется

$$KC(\mathfrak{M}_D) = \min_{U \notin \mathscr{U}_s} \max_{\mathscr{M}_D \in \mathfrak{M}_D} KC_U(\mathscr{M}_D).$$

Важно заметить, что в определении точной колмогоровской сложности семейства моделей берётся «оптимальная» по сжатию самой сложной модели в классе $\mathfrak{M}_D$ машина Тьюринга. Если существует более одной такой МТ (тогда для всех из них значение $KC(\mathfrak{M}_D)$ одинаково), то берется одна любая «оптимальная» по сжатию МТ.

Определение 6. *Точной колмогоровской сложностью семейства* $\tilde{\Omega}_D$ согласованных систем линейных ограничений называется

$$KC(\tilde{\Omega}_D) = \min_{U \notin \mathscr{U}_s} \max_{\Omega_D \in \tilde{\Omega}_D} KC_U(\Omega_D).$$

Точная колмогоровская сложность не является вычислимой функцией. Для использования этого понятия будут применяться оценки сложности сверху. С техникой получения подобных оценок можно ознакомиться в [5].

4. Обоснование согласованных моделей на основе колмогоровской сложности

Описанием числовых векторов и матриц являются последовательности подходящим образом закодированных чисел. Если значительная часть коэффициентов матрицы (векторов) согласованной модели — нулевые, то описание (сложность) модели может оказаться существенно короче.

Лемма 1. *Согласованные модели из семейства* $\mathfrak{N}_D$*, имеющего точную колмогоровскую сложность* d*, могут быть извлечены путём применения «оптимальной» по сжатию машиной* U_D *не более чем* 2^d *способами.*

Доказательство. Пусть для решения поставленной задачи синтеза применяется произвольное согласованное семейство $\mathfrak{N}_D \subseteq \mathfrak{M}_D$ такое, что $KC(\mathfrak{N}_D) = d$, и зафиксирована некоторая «оптимальная» по сжатию МТ U_D. Применение этой машины к двоичной строке $p : |p| = d$ обеспечивает получение линейной модели $\mathscr{M}_D(p)$, представленной в форме конечной строки-описания $S_p = U_D(p)$, где $S_p = s_{\hat{C}} {*} s_{\hat{A}} {*} s_{\hat{B}}$. Поскольку U_D является функцией ($U_D \in P_{pr}$), каждому значению p соответствует единственная строка $S_p = U_D(p)$. Тогда 2^d различных возможных значений строки p длины d исчерпывают все порождаемые строки вида $S_p = s_{\hat{C}} * s_{\hat{A}} * s_{\hat{B}}$ и, возможно, строки, не имеющие смысла как, например, код согласованной модели $S_p = U_D(11...11)$. $\square$

Следствие 1. *Пусть* $KC(\mathfrak{N}_D) < h$*, где* h *— верхняя оценка сложности семейства* $\mathfrak{N}_D$*. Тогда по кодам (программам) длины* h *можно получить не более* 2^h *согласованных моделей, и* $|\mathfrak{N}_D| \leqslant 2^h$*.*

Напомним, что изучаемая задача синтеза будет решаться на вычислительной машине, в которой число бит, используемое для представления одного числа, равно b; именного столько бит будет использоваться для записи одного числа и в обучающих данных D, и в описании кода модели $S_p = s_{\hat{C}} * s_{\hat{A}} * s_{\hat{B}}$. Тогда в семействе $\mathfrak{M}_D$ будет конечное число моделей. И программирование верхних оценок сложности [5] будет связано с конструированием строк, длина которых будет определяться, в числе прочего, константой b.

Обозначим $conv(D_\oplus)$ выпуклую оболочку множества допустимых точек из обучающей выборки.

Лемма 2. *При условии $D_\ominus \cap conv(D_\oplus) = \varnothing$, мощность семейства $|\tilde{\Omega}_D|$ удовлетворяет неравенству*

$$2^{log(\lceil \hat{\rho}/\delta \rceil - 2) l_\ominus} \leqslant |\tilde{\Omega}_D|,$$

где $\hat{\rho}$ — наименьшее по всем точкам обучающего множества $D_\ominus$, лежащим вне множества допустимых решений, евклидово расстояние до выпуклой оболочки точек из $D_\oplus$, которые являются допустимыми решениями, а δ — модуль наименьшего рационального числа в используемом машинном формате.

Доказательство. Пусть $\hat{\rho} = \min_{X_k \in D_\ominus} \rho(X_k, conv(D_\oplus))$ — кратчайшее расстояние по всем точкам множества $D_\ominus$ до выпуклой оболочки точек множества $D_\oplus$. Отрезок прямой, длина которой равна $\hat{\rho}$, соединяющий точку из $D_\ominus$ с точкой оболочки, можно отделить от $conv(D_\oplus)$ не менее чем $\lceil \hat{\rho}/\delta \rceil - 2$ перпендикулярными ему гиперплоскостями. Поэтому каждая точка из $D_\ominus$ линейно отделима не менее чем $\lceil \hat{\rho}/\delta \rceil - 2$ гиперплоскостями, и из них можно выбрать одну любую для отделения каждой точки. Тогда число линейных моделей, отделяющих все точки множества $D_\ominus$, не менее $(\lceil \hat{\rho}/\delta \rceil - 2)^{l_\ominus} = 2^{log(\lceil\ \hat{\rho}/\delta \rceil - 2) l_\ominus}$. □

Решение задачи поиска модели, «объясняющей» выборочные данные, происходит как машинное обучение, приводящее к извлечению гипотезы — закономерности в данных. Важно оценить неслучайность выбора гипотезы.

В теории статистических решений предполагается существование вероятностных распределений в пространстве выбора гипотез. В данной работе рассматривается нестатистическая постановка задачи: выборочные данные отражают регулярные модели. Поэтому применяется подход к выбору модели в условиях неопределенности, предполагающий возможность равновероятного извлечения любой согласованной модели. Но синтез согласованной модели, как процесс машинного обучения, должен приводить не к случайной гипотезе, а к закономерности.

Для понимания доказательства следующих утверждений нужно обратить внимание на то, что сложность линейной согласованной модели всегда не меньше сложности входящей в неё системы ограничений.

Теорема 1. *Пусть модели согласованного семейства* $\mathfrak{M}_D$ *распределены в нем равновероятно. Тогда вероятность случайного извлечения согласованной модели из класса* $\mathfrak{N}_D \subset \mathfrak{M}_D$*, имеющего оценку точной колмогоровской сложности* $KC(\mathfrak{N}_D) < h$*, не превысит*

$$2^{h-log(\lceil\hat{\rho}/\delta\rceil-2)l_\ominus}$$

Доказательство. Напомним, что семейство $\mathfrak{M}_D$ является конечным в силу условия сужения моделей до реализуемых на компьютере с ограничением рационального промежутка представления чисел. Оцениваемая вероятность случайного выбора есть $|\mathfrak{N}_D|/|\mathfrak{M}_D|$. Согласно условию теоремы, $KC(\mathfrak{N}_D) < h$, и $|\mathfrak{N}_D| \leqslant 2^h$ по следствию 1. Согласно лемме 2,

$$|\mathfrak{M}_D| \geqslant 2^{log(\lceil\hat{\rho}/\delta\rceil-2)l_\ominus},$$

поскольку модели класса $\mathfrak{M}_D$ включают в себя все модели линейных ограничений семейства $\tilde{\Omega}_D$. Поэтому

$$|\mathfrak{N}_D|/|\mathfrak{M}_D| \leqslant 2^h/|\mathfrak{M}_D| \leqslant 2^h/2^{log(\lceil\ \hat{\rho}/\delta\rceil-2)l_\ominus} = 2^{h-log(\lceil\hat{\rho}/\delta\rceil-2)l_\ominus}$$

□

Замечание. Полученная в теореме 1 оценка справедлива для расширения компьютерных моделей до бесконечного класса вещественных.

Следствие 2. *Для того, чтобы построенная модель* $\mathscr{M}_D$*, имеющая верхнюю оценку точной колмогоровской сложности* h*, с вероятностью не меньшей* $1-\varepsilon$ *была неслучайной, т. е. была закономерностью, необходимо, чтобы число примеров в выборке удовлетворяло неравенству*

$$l_\ominus \geqslant (h + log(1/\varepsilon))/log(\lceil\hat{\rho}/\delta\rceil - 2). \quad (3)$$

Доказательство. Основываясь на колмогоровском представлении о достоверности как неслучайности, из неравенства

$$2^{h-l_\ominus log(\lceil\hat{\rho}/\delta\rceil-2)} \leqslant \varepsilon$$

легко вывести эквивалентное неравенство (3). □

Например, $l_\ominus(\varepsilon = 0.01,\ \hat{\rho} = 1,\ \delta = 10^{-4}, h = 3000) \approx 230$.

Согласно полученным результатам, необходимая величина объема $l_\ominus$ подвыборки $D_\ominus$ определяется, главным образом, колмогоровской сложностью модели (оценкой сложности) и кратчайшим расстоянием $\hat{\rho}$, взятым по всем точкам множества $D_{\ominus}$, до выпуклой оболочки точек множества $D_{\oplus}$. Невхождение длины $l_{\oplus}$ подвыборки в явном виде в оценку (3) объясняется тем, что от нее в значительной степени зависит оценка кратчайшего расстояния $\hat{\rho}$, которая в эту оценку входит явно. Рост числа «положительных» примеров $l_{\oplus}$ в обучающей информации влечет всё более равномерное заполнение области допустимых решений и уменьшение $\hat{\rho}$.

VC-размерность Вапника-Червоненкиса [4] $h_{VC}(\mathfrak{M})$ некоторого семейства моделей $\mathfrak{M}$ связана с его колмогоровской сложностью $\mathrm{KC}(\mathfrak{M})$ соотношением

$$h_{VC}(\mathfrak{M}) \leqslant \mathrm{KC}(\mathfrak{M}) < h_{VC}(\mathfrak{M}) logl,$$

где l — длина обучающей выборки [17]. Это показывает, что нестатистическая оценка качества моделей — закономерностей хорошо согласуется со статистической теорией В. Н. Вапника и А. Я Червоненкиса.

5. Рекомендации к построению согласованных моделей

1. Необходимо стремиться к получению модели с минимальной колмогоровской сложностью. Достигнуть снижение сложности синтезированной модели можно за счет:
   – минимизации числа неравенств в $\hat{\Omega}_D$;
   – минимизации ненулевых членов оценочных матриц и векторов, поскольку pVCD оценка [5] сжатия строки описания модели уменьшается при этом на число нулевых элементов, умноженное на константу b.

   При построении неравенств согласованной линейной оптимизационной модели каждая точка выборочных данных из $D_{\ominus}$ должна линейно отделяться ото всех допустимых точек из $D_{\oplus}$. Поэтому всегда можно синтезировать $l_{\ominus}$ гиперплоскостей. Затем следует произвести переобучение для каждой из них, включив в выборку для каждой плоскости уже все отделённые ею точки. И только затем решать задачу о кратчайшем наборе гиперплоскостей, как задачу минимизации [8].

2. Алгоритмы построения линейных отделителей необходимо усовершенствовать введением выполняемого в ходе синтеза «сброса» в ноль коэффициентов, значение которых ниже некоторого порогового значения. Такая модификация может позволить понизить сложность линейных функций, для синтеза которых представляется полезным применение алгоритмов SVM [14].

3. Необходимо тщательно разработать структуру слова p, являющееся сжатым описанием модели. Уменьшение его длины может быть достигнуто, например, за счет экономного кодирования идущих «подряд» нулей в описании модели. Это позволит уменьшить значение верхней оценки сложности модели h.

4. Если в n-мерном пространстве задано $l_{\oplus}$ точек то при $n \geqslant 6$ число граней выпуклой оболочки имеет порядок $O({l_{\oplus}}^{\lfloor n/2 \rfloor})$ [13]. Следовательно, построение выпуклой оболочки точек множества $D_{\oplus}$ трудоёмко. Указанная в работе [3] процедура $FullConvexhull$ строит описание граней всех размерностей оболочки и нормальных конусов к ним. Подобные процедуры слишком сложны в случае большой размерности. Для оценивания кратчайшего расстояния до выпуклой оболочки её построение в явном виде не требуется.

Заключение

В статье предложен колмогоровский подход к определению сложности линейных согласованных моделей оптимизации. На основе оценивания колмогоровской сложности моделей даны рекомендации к обоснованию методов синтеза таких моделей по прецедентной начальной информации.

Вопросы формирования обучающей прецедентной информации в статье не затронуты. Но растущая большими темпами компьютеризация объектов и большие объемы доступной для хранения информации памяти на сегодняшний день позволяют извлекать из хранилищ данных выборки, имеющие длину свыше десятков тысяч наблюдений.

Направление дальнейшей работы связано с усовершенствованием алгоритмов синтеза линейных согласованных моделей с учетом данных выше рекомендаций.

Список литературы

1. Анафиев А. С. М-Модели алгоритмов. Ёмкость и колмогоровская сложность класса М-полиномов / А. С. Анафиев // Таврический вестник информатики и математики. – 2010. – № 1. – С. 51–57.
2. Блыщик В. Ф. Интеллектуализированная программная система Intman поддержки принятия решений в задачах планирования и управления / В. Ф. Блыщик, В. И. Донской, Г. А. Махина // Искусственный интеллект. — 2002. — С. 406–415.
3. Буровский П. А. Алгоритм Моцкина-Бургера и вычисление выпуклых оболочек точек n - мерного пространства / П. А. Буровский // Вестник КрасГУ. Сер. физ.-мат. науки. — Красноярск, 2005. — Вып.1. — С. 48–53.
4. Вапник В. Н. Теория распознавания образов / В. Н. Вапник, А. Я. Червоненкис. — М.: Наука, 1974. — 416 с.

5. Донской В. И. Оценки емкости основных классов алгоритмов эмпирического обобщения, полученные pVCD методом / В. И. Донской // Ученые записки ТНУ им. В. И. Вернадского. Серия «Физико-математические науки», 2010. — Т.23 (62). — №2. — С. 56–65.
6. Донской В. И. Слабоопределенные задачи линейного булева программирования с частично заданным множеством допустимых решений / В. И. Донской // ЖВМ и МФ, 1988. — Т. 28. — №9. — С. 1379–1385.
7. Донской В. И. Сложность семейств алгоритмов обучения и оценивание неслучайности извлечения эмпирических закономерностей / В. И. Донской // Кибернетика и системный анализ. — 2012. — №2. — С. 86–96.
8. Донской В. И. Частично определенные задачи оптимизации: подход к решению на основе теории распознавания образов / В. И. Донской // Динамические системы. — К.: Вища школа, 1989. — Вып. 8. — С. 71–77.
9. Колмогоров А. Н. Теория информации и теория алгоритмов / А. Н. Колмогоров — М.: Наука, 1987. —- 304 с.
10. Мазуров Вл. Д. Применение методов теории распознавания образов в оптимальном планировании и управлении Вл. Д. Мазуров // Труды I Всесоюзной конференции по оптимальному планированию и управлению народным хозяйством. — М., ЦЭМИ, 1971. — С. 49.
11. Мазуров Вл. Д. Нестационарные процессы математического программирования. — М.: Наука, 1979. — 288 с.
12. Мазуров Вл. Д. Распознавание образов как средство автоматического выбора процедуры в вычислительных методах / Вл. Д. Мазуров // ЖВМ и МФ, 1970. — Т. 10. — №6. — С. 1520—1525.
13. Препарата Ф. Вычислительная геометрия. Введение / Ф. Препарата, М. Шеймос. — М.:Мир, 1989. — 478 с.
14. Таратынова Н. Ю. Задача линейной оптимизации с частично заданной информацией / Н. Ю. Таратынова // Таврический вестник математики и информатики, 2005. — №1. — С. 82–93.
15. Таратынова Н. Ю. Построение оптимизационной модели по прецедентной начальной информации как задача нелинейной регрессии / Н. Ю. Таратынова // Искусственный интеллект, 2006. — №2. — С. 82–93.
16. Chan T. M. Output-Sensitive Construction of Convex Hulls / T. M. Chan / Ph.D. thesis. — Department of Computer Science, University of British Columbia. — 1995. — 104 p.
17. Donskoy V. I. The Estimations Based on the Kolmogorov Complexity and Machine Learning from Examples / V. I. Donskoy // Proceedings of the Fifth International Conference „Neural Networks and Artificial Intelligence" (ICNNAI'2008). — Minsk: INNS. — 2008. — P. 292–297.
18. Donskoy V. I. Pseudo-Boolean Scalar Optimization Models with Incomplete Information / V. I. Donskoy. — GMOOR Newsletters, 1996. — № 1–2. — P. 20–26.
19. Li Ling. Data Complexity in Machine Learning and Novel Classification Algorithms /Ling Li // Thesis for the Degree of PhD. — California Institute of Technology, 2006. — 103 p.

Статья поступила в редакцию 27.05.2012