



УЛЬЯНОВСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА

Ученые записки

экология

экономические науки

клиническая медицина

фундаментальные проблемы математики и механики

гуманитарные науки и механики

биология и медицина

лингвистика

государство и право



Министерство образования Российской Федерации

Ульяновский государственный университет

УЧЕНЫЕ ЗАПИСКИ
Ульяновского государственного университета

Серия

***ФУНДАМЕНТАЛЬНЫЕ ПРОБЛЕМЫ
МАТЕМАТИКИ И МЕХАНИКИ***

Выпуск 2(9)

Ульяновск 2000

ББК 22.1+22.2
У 91

*Печатается по решению
ученого совета механико-математического факультета
Ульяновского государственного университета*

**У 91 Ученые записки Ульяновского государственного
университета. Сер. Фундаментальные проблемы мате-
матики и механики. Вып. 2(9) / Под ред. академ. РА-
ЕН, проф. А.С.Андреева. - Ульяновск: УлГУ, 2000. -
113 с.**

В сборнике публикуются статьи преподавателей и аспирантов механико-математического факультета по фундаментальным проблемам математики и механики.

Сборник представляет интерес для научных работников, аспирантов и студентов, занимающихся вопросами алгебры и математической логики, математической кибернетики и теории вероятности, информационных технологий и программирования, теоретической механики и механики твердого тела.

© Ульяновский государственный
университет, 2000

МОДЕЛИРОВАНИЕ ВЕРОЯТНОСТИ ОБРАЗОВАНИЯ СРЕЗА С УЧЕТОМ РАДИУСОВ ЗАКРУГЛЕНИЯ ЭЛЛИпсоИДАЛЬНЫХ ЗЕРЕН ШЛИФОВАЛЬНОГО КРУГА И ИХ ОРИЕНТАЦИИ¹

А.Ю. Богданов, В.В. Богданов

Способность материала заготовки полностью или частично восстанавливать свою форму после снятия нагрузки является неотъемлемой частью процесса механической обработки, в частности, абразивной обработки. Предпринимались попытки учесть в расчетах вероятности образования среза единичным зерном физико-механические свойства материала шлифуемой заготовки, то есть существование упругой деформации материала. Особенностью предлагаемого подхода является представление абразивных зерен в форме эллипсоидов вращения, что позволит моделировать различную ориентацию зерен в пространстве, различные передние углы зерен, а также износ абразивного зерна (за счет изменения переднего угла) во времени (рис. 1).

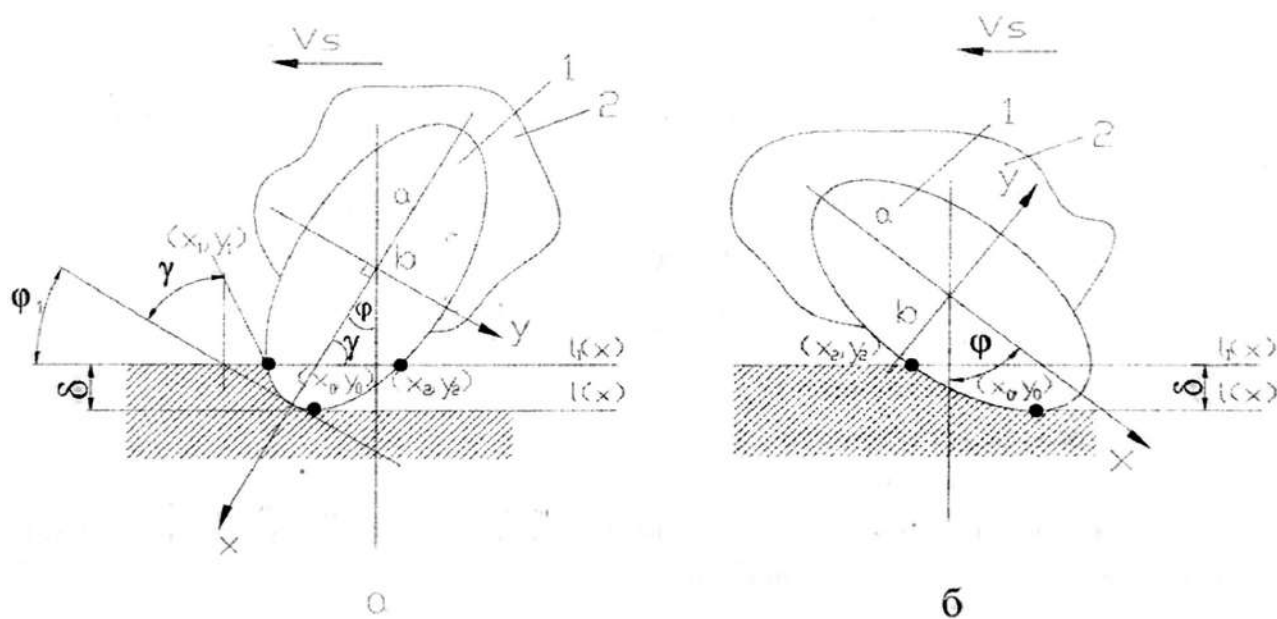


Рис.1. Плоская схема резания зерном в виде эллипса:

1 - зерно; 2 - связка; а) схема 1; б) схема 2

Пусть полуоси эллипсоида связаны зависимостью $b=ka$, где $k \in [0; 1]$, случайная величина с известной плотностью распределения $f_k(x)$ (можно в первом приближении считать равномерным, то есть $f_k(x) \equiv 1$).

Распределение абразивных частиц инструмента описывается распределением по диаметрам d_0 (частицы считаются шарами), поэтому для дальнейших рассуждений потребуются соотношения: объем шара $V_{ш} = \pi d_0^3/6$; объем эллипсоида вращения $V_э = 4/3$

¹ Работа поддержана Российским фондом фундаментальных исследований, проект N 00-01-00312.

πab^2 . После приравнивания объемов полуоси эллипсоида a и b выражаются через диаметр частицы d_0

$$a = d_0 / (2k^{2/3}), \quad b = ka = d_0 k^{1/3} / 2.$$

Данный подход позволяет учитывать все разнообразие абразивных зерен круга, так как допускается существование шарообразных и игольчатых зерен.

Эллипсоидальное зерно в связке может быть закреплено по одной из двух схем (рис. 1), отличающихся углом φ между большой осью зерна и нормалью к поверхности заготовки. Таким образом $\varphi \in [-\pi/2; \pi/2]$ - угол ориентации главной оси эллипсоидального зерна. Это случайная величина, у которой будем считать известной плотность $f_\varphi(x)$. Для определенности распределение можно считать равномерным.

Авторами построена конструктивная процедура нахождения координат точек $(x_0; y_0)$, $(x_1; y_1)$ и $(x_2; y_2)$ при известных d_0 , φ , k , δ , которая ввиду громоздкости выражений здесь не приводится. В терминах d_0 и k радиусы закругления в точках $(x_i; y_i)$ равны

$$\rho_i = \frac{4}{k^{8/3} d_0^2} (x_i^2 k^4 + y_i^2)^{3/2}, \quad i = 0, 1, 2.$$

Введем в рассмотрение средние значения

$$\rho_{cp} = (\rho_0 + \rho_1) / 2, \quad \text{и} \quad \rho_{cp} = (\rho_0 + \rho_2) / 2,$$

которые наиболее точно соответствует радиусу закругления режущей кромки абразивного зерна при $\varphi \in [0; \pi/2]$ для схемы 1 (рис. 1, а), и при $\varphi \in [-\pi/2; 0]$ для схемы 2 (рис. 1, б), соответственно.

Функция возможности осуществления среза будет иметь вид

$$R(d_0, \varphi, k, \delta) = \begin{cases} 0, & \text{если } \delta < (1 - \tau_s / \sigma_t) \rho_{cp} \\ 1, & \text{если } \delta \geq (1 - \tau_s / \sigma_t) \rho_{cp} \end{cases},$$

где τ_s - предел прочности материала на сдвиг;

σ_t - предел текучести материала;

δ - толщина среза.

Выражение для вероятности образования среза единичным зерном при плоском шлифовании на глубине x будет иметь вид

$$P^1(x) = \int_0^{+\infty} f_a(d_0) \int_{-\pi/2}^{\pi/2} f_\varphi(\varphi) \int_0^1 f_k(k) \int_x^h R(d_0, \varphi, k, y-x) \frac{\sqrt{2(R_k - h + y)(y-x) - (y-x)^2}}{\sqrt{2R_k h - h^2}} f_B^a(y+a) dy dk d\varphi dd_0$$

где $f_a(d_0)$ - плотность вероятности распределения абразивных зерен по диаметрам,

$f_B^a(x)$ - динамическая плотность распределения вершин зерен абразивного круга [1],

R_k - радиус круга, h - активная распределенная кромка круга.

Существование промежуточных состояний расположения зерна между схемами 1 и 2 позволяет моделировать не только различное расположение переднего угла абразивного зерна, но и предполагает возможность исследования динамики абразивных зерен в процессе шлифования. Тогда в полученных зависимостях необходимо ввести составляющую времени, что приведет к вычислению еще одного интеграла.

Литература

1. Богданов В.В., Богданов А.Ю. Динамический закон распределения вершин зерен абразивного инструмента с учетом его механического износа и загрязнения. Ученые записки Ульяновского государственного университета. Серия "Фундаментальные проблемы математики и механики". Выпуск 1(6). 1999. С. 56-64.

ОБ УСЛОВИЯХ КОММУТИРОВАНИЯ В ГЛОБАЛЬНОМ НАДМОНОИДЕ СВОБОДНОГО МОНОИДА

А.Г. Бросалина

Данная статья – попытка обобщения несложного утверждения об условиях коммутирования слов на случай языков. В статье приведены полученные необходимые условия для коммутирования языков, при этом в префиксном случае получены необходимые и достаточные условия. Важно отметить, что рассматриваемые алфавиты и языки могут быть бесконечными.

Показаны возможности применения полученных результатов в других задачах теории формальных языков. Сформулированы связанные задачи для дальнейшего решения. **Введение**

Во многих учебниках по теории формальных языков (см.напр. [1]) рассматривается следующее несложное утверждение.

Утверждение 1. *$uv = vu$ тогда и только тогда, если*

$$(\exists w \in \Sigma^*, k, l \in \mathbb{N}_0) (u = w^k, v = w^l).$$

□

Данная статья – попытка обобщения данного утверждения об условиях коммутирования слов на случай языков. До конца рассматриваемая проблема не решена – т.е. *в общем случае* удобных для использования в других задачах необходимых и достаточных условий этого равенства не получено ¹, и ещё не решенные задачи сформулированы в виде гипотез. В настоящей статье приведены полученные необходимые условия, а в префиксном случае – необходимые и достаточные.

В предыдущих публикациях одного из авторов (см.напр. [2]) частные случаи данных утверждений были применены при решении эквивалентных проблем в специальных подклассах класса контекстно-свободных языков. В настоящей статье получены и более сложные утверждения – в частности, когда рассматриваемые алфавиты (или языки над ними) являются бесконечными.

Авторы надеются, что критерии, полученные в этой работе для более общих случаев, могут быть использованы в некоторых других задачах теории формальных

¹ Такими “удобными для использования” можно считать, например, алгоритмически разрешимые задачи.

языков. (Некоторые из таких задач были рассмотрены в [2]. В этой же работе была сформулирована тематика для дальнейшего исследования, т.е. комплекс связанных задач.)

Предварительные замечания

Рассмотрим свободный моноид $(\Sigma, \cdot, e)$ и его подмоноид $(\mathcal{P}(\Sigma), \cdot, \{e\})$.

Некоторые определения, связанные с теорией формальных языков, приведены в [1, 3]. Введем дополнительно несколько определений.

Пусть $v = a_1 a_2 \dots a_n$ и для некоторых $m \leq n$ мы имеем $u = a_1 a_2 \dots a_m$, тогда u называется префиксом слова v . Мы будем обозначать этот факт как $u \in \text{pref}(v)$.

Пусть U и V удовлетворяют условию:

$$(\forall u \in U)(\exists v \in U)(u \in \text{pref}(v)).$$

Тогда если $U \in \text{pref}(V)$, то будем писать $U \subseteq V$.

Отметим, что ниже мы будем обозначать бесконечные языки большими прописными буквами $(A, \dots, D)$, а конечные — $A, \dots, D$.

Для произвольного языка A обозначим

$$\text{pv}(A) = \{u \in A \mid (\forall v \in A)(u \notin \text{pref}(v))\}.$$

Например, пусть над алфавитом $\Delta = \{0, 1\}$ заданы языки

$$C = \{0, 100, 101, 111\}, D = C \cup \{01, 110\}.$$

Тогда для этих языков имеем

$$\text{pv}(D) = C.$$

Обозначим длину самого короткого слова языка A как $\|A\|_{\min}$.

Если $(\forall u, u \in A)(u \notin \text{pref}(v))$, то будем называть множество (язык) A префиксным, записывая этот факт $\text{Pr}(A)$.

Далее введем определение максимального префиксного кода.

Кодом над алфавитом A называется всякое подмножество из A^* , являющееся базисом свободного подмоноида из A^* . Т.е. $C \subseteq A^*$ есть код, если произвольное равенство

$$x_1 \dots x_m = y_1 \dots y_n,$$

где

$$x_1, \dots, x_m, y_1, \dots, y_n \in C,$$

влечет

$$m = n \text{ и } x_1 = y_1, \dots, x_n = y_n.$$

Код $C \subseteq A^*$ называется префиксным, если никакое слово из C не является левым делителем другого слова из C , т.е. $CA^* \cup C = \emptyset$.

Код над алфавитом A называется максимальным, если он не содержится ни в каком другом коде над A .

Теперь для произвольного алфавита Δ рассмотрим следующее рекуррентное определение множества максимальных префиксных кодов над Δ . Будем обозначать определяемое множество языков $\text{mp}(\Delta)$.

- Считаем, что $\Delta \in \text{mp}(\Delta)$.
- Для произвольных $C \in \text{mp}(\Delta)$ и $B \in C$ полагаем

$$(C \setminus B \cup B\Delta \in \text{mp}(\Delta)).$$

Если множество D таково, что для некоторого $C \in \text{mp}(\Delta)$ выполнено включение $C \subseteq D$, то будем писать $D \in \text{mp}^+(\Delta)$.

Рассмотрим примеры к нескольким последним определениям. Пусть над алфавитом $\Delta = \{0, 1\}$ заданы языки

$$C = \{0, 100, 101, 111\}, D = C \cup \{01, 110\}.$$

Тогда

$$C \in \text{mp}(\Delta),$$

кроме того, $C, D \in \text{mp}^+(\Delta)$.

Для алфавита $\Sigma = \{a, b\}$ и множества $A = \{ab, bba\}$ можно, например, считать, что $\Delta_A = \Delta$, а морфизм

$$h_A : \Delta_A^* \rightarrow \Sigma^*$$

задан таким образом:

$$h_A(0) = ab, h_A(1) = bba.$$

Тогда выполняется следующее:

$$h_A(C) = \{ab, bba ab ab, bba ab bba, bba bba\} \in \text{mp}(A),$$

$$h_A(D) = \{ab, bba ab ab, bba ab bba, bba bba, ab bba, bba bba ab\} \in \text{mp}^+(A).$$

Итак, будем рассматривать равенство

$$AB = BA \tag{1}$$

на множествах языков (т.е. – в глобальных надмоноидах свободных моноидов).

Отметим, что рассматриваемые множества A и B могут содержать e – именно из-за этого появляются проблемы, рассматривающиеся в разделе . Более того, если e не

содержится ни в одном из рассматриваемых множеств, то, например, доказываемое далее утверждение 2 тривиально.

Случай произвольных конечных языков

Итак, сначала рассмотрим случай конечных языков. Без доказательства отметим следующий очевидный факт: если $AB = BA$, то

$$(\forall k, l \in \mathbb{N}_0) (A^k B^l = B^l A^k). \quad 2$$

Утверждение 2. Если $AB = BA$, то $B^* \subseteq A^*$.

Доказательство. Как было сказано ранее, в этом разделе рассмотрим только случай конечных множеств A и B .³

Предположим сначала, что $A \not\ni e$. Рассмотрим любое слово $v \in B^*$, выберем $l \in \mathbb{N}_0$, для которого $v \in B^l$; обозначим $k = |v| + 1$. Выполнено условие

$$v \in \text{opref}(B^l A^k)$$

($e \notin A$, значит и $e \notin A^k$, поэтому в последнем случае действительно можно писать “opref”, а не только “pref”). Согласно отмеченному выше, $B^l A^k = A^k B^l$, поэтому

$$v \in \text{opref}(A^k B^l).$$

Поскольку $e \notin A$, имеем

$$\|A^*\|_{\min} \geq k > |v| \quad \text{и} \quad v \in \text{opref}(A^k).$$

Нами было рассмотрено произвольное слово $v \in B^*$, значит при $e \notin A$ условие

$$B^* \subseteq \text{opref}(A^*)$$

(т.е. $B^* \subseteq A^*$) выполнено. Далее в доказываемом утверждении считаем, что $A \ni e$.

Предположим, что $B^* \not\subseteq \text{opref}(A^*)$, т.е.

$$(\exists v \in B^*) (v \notin \text{opref}(A^*)).$$

Выберем $l \in \mathbb{N}$ таким образом, чтобы $v \in B^l$, и рассмотрим язык

$$\mathcal{D} = \{w \in A^* \mid w \in \text{pref}(v)\}.$$

Заметим, что

$$v \notin \text{opref}(A^*),$$

² И вообще, все конкатенации из $k + l$ языков, среди которых k раз встречается A и l раз — B , равны между собой.

³ Следовательно, конечным можно считать и рассматриваемый алфавит Σ .

значит $v \notin D$, поэтому можно альтернативным способом определить $\mathcal{D}$ как

$$\mathcal{D} = \{w \in A^* \mid w \in \text{orgref}(v)\}.$$

Выберем некоторое слово $u \in \max(\mathcal{D})$, пусть

$$u \in A^{k_1} \quad \text{и} \quad |u| = m$$

(выберем таким образом числа k_1 и m). Поскольку $A \neq \{e\}$, то при любых $i < j$ имеем место неравенство

$$|A^i| < |A^j|.$$

Поэтому можно выбрать некоторое $k_2 \in \mathbb{N}$, такое что

$$|A^{k_2}| > (m+1) \cdot |B^l|.$$

Обозначим

$$k = \max(k_1, k_2).$$

Заметим, что согласно условиям $e \in A$ и $u \in A^{k_1}$, слово u входит в каждый из языков A^i при $i \geq k_1$, значит и $A^k \ni u$.

Обозначим

$$v = ua_1a_2 \dots a_n,$$

где

$$a_1, a_2, \dots, a_n \in \Sigma,$$

и, очевидно, $n > 0$; определим языки

$$C_1 = \{u, ua_1, ua_1a_2, \dots, ua_1a_2 \dots a_{n-1}\},$$

$$C_2 = v\Sigma^*, \quad C = C_1 \cup C_2.$$

Рассмотрим язык $D = A^k B^l$. Согласно отмеченному выше, верно также равенство $D = B^l A^k$. Оценим двумя способами – сверху и снизу – число (различных) элементов множества $D \cap C$.

Во-первых, рассмотрим запись D в виде $A^k B^l$. Все слова вида uw , где $w \in B^l$, могут принадлежать $D \cap C$. Поскольку, согласно предположению $|u| = m$, слово u имеет ровно m различных собственных префиксов (включая e), а включение

$$u'w \in D \cap C$$

для некоторого $w \in B^l$ может быть выполнено для произвольного слова u' последнего множества ($u' \in \text{orgref}(u)$), то общее число элементов множества $D \cap C$ оценивается *сверху* следующим образом:

$$|D \cap C| \leq |B^l| + m \cdot |B^l| = (m+1) \cdot |B^l|. \quad (3)$$

(Мы действительно посчитали все слова множества $D \cap C$, т.к. условие $u'' \in C$ при $u'' \in A^k$ и $u'' \neq u$ не может быть выполнено по следующей причине. Если бы u'' входило в язык C_1 , то был бы неверен способ выбора u как элемента множества $\text{max}(\mathcal{D})$, поскольку

$$u'' \in A^* \quad \text{и} \quad |u''| > |u|.$$

А если бы выполнялось условие $u'' \in C_2$, то выполнялось бы также и включение $v \in \text{opref}(A^*)$, а последнее противоречит способу выбора v .)

Итак, не существует никаких элементов языка $D \cap C$, кроме пересчитанных, т.е. выполнено неравенство (3).

Во-вторых, рассмотрим запись D в виде $B^l A^k$. Вследствие $A \in e_2$ выполнено условие $v \in B^l$, поэтому все элементы множества D вида vw , где $w \in A^k$, принадлежат также C_2 , и, следовательно, C . Таким образом, число (различных) элементов языка $D \cap C$ не меньше числа различных элементов A^k , а согласно условию $k \geq k_2$ и способу выбора k_2 , имеем

$$|D \cap C| > (m+1) \cdot |B^l|. \quad (4)$$

Итак, получено противоречие – (3) и (4), доказывающее условие

$$B^* \subseteq \text{opref}(A^*),$$

т.е. $B^* \infty A^*$. □

С помощью двукратного применения утверждения 2 доказывается

Теорема 1. Если $AB = BA$, то $A^* \infty B^*$. □

Применим к каким-либо языкам A и B , удовлетворяющим условию (1), теорию из статьи [2]. Получаем, что из исследуемого равенства (1) и его следствия – условия $A^* \infty B^*$ (теорема 1) – вытекает, что над алфавитом Σ^* существует некоторое множество C , такое что

$$A, B \in \text{mp}^+(C).$$

Или в альтернативной формулировке:

$$\exists h_C^{-1}(A), h_C^{-1}(B) \in \text{mp}^+(\Delta_C).$$

Заметим, что хотя для некоторого $u \in A \cup B$ множество $h_C^{-1}(u)$ может при этом содержать более одного элемента (т.к. морфизм h_C^{-1} , соответствующий языку C , не обязательно является инъективным), однако языки

$$A' = h_C^{-1}(A) \quad \text{и} \quad B' = h_C^{-1}(B)$$

удовлетворяют соотношению

$$A' B' = B' A' \quad (5)$$

Доказательство утверждения (5) проводится от противного с помощью применения к рассматриваемым множествам морфизма h_C .

Итак, мы можем применять ко множествам A и B , удовлетворяющим равенству (1), специальный инверсный морфизм, получая аналогичное соотношение (5) для новых, "раскодированных" множеств A' и B' . Более того, известно, что каждый из языков A' и B' содержит в качестве подмножества какой-нибудь максимальный префиксный код над алфавитом Δ_C . Некоторые возможные алгоритмы выбора подходящего множества C были обсуждены в [2, 4].

И, как было сказано ранее, некоторые из задач теории формальных языков, возможных для применения полученных в данном разделе критериев, были рассмотрены в [2]. Однако в тех работах существенным является то, что рассматриваемые языки могут быть бесконечными. Поэтому в следующем разделе рассмотрим обобщение утверждений, доказанных в этом разделе.

Случай произвольных бесконечных языков

Утверждения, рассмотренные в этом разделе, являются обобщением теоремы, доказанной для множеств, которые имеют конечное число атомов. Теперь докажем их для случая бесконечных множеств.

Как уже говорилось ранее, множества, обозначенные A, B (и т.д.), могут содержать бесконечно много слов.

Обозначим подязык, который состоит из слов языка A , имеющих длину i как A_i .

Для множеств бесконечных языков над алфавитом Σ определим специальный частичный линейный порядок.

Это может быть обозначено как обобщение простого отношения "содержит больше слов".

Определение 1. *Определим линейный порядок $\succ$ для конечных и бесконечных алфавитов*

- Пусть алфавит Σ является конечным.

Мы будем писать $A \succ B$, если существует n , такое, что выполняются условия:

$$|A_1| = |B_1|, |A_2| = |B_2|, \dots, |A_{n-1}| = |B_{n-1}|, |A_n| > |B_n|.$$

- Пусть алфавит Σ является бесконечным.

Мы будем писать $A \succ B$, если существует такое n , что выполняются следующие условия:

$$\begin{aligned} A_1 &= B_1 \quad \text{или} \quad (A_1 \not\subset B_1 \quad \text{и} \quad B_1 \not\subset A_1), \\ A_2 &= B_2 \quad \text{или} \quad (A_2 \not\subset B_2 \quad \text{и} \quad B_2 \not\subset A_2), \\ &\dots \quad \dots \quad \dots \\ A_{n-1} &= B_{n-1} \quad \text{или} \quad (A_{n-1} \not\subset B_{n-1} \quad \text{и} \quad B_{n-1} \not\subset A_{n-1}), \\ A_n &\supset B_n. \end{aligned}$$

Например, для произвольного алфавита Σ в языке Σ^* "содержится больше слов", чем в языке $(\Sigma^2)^*$. Очевидно, что если $A_n \subset B_n$, то $A_n \prec B_n$.

Теперь докажем факт, который является обобщением теоремы 2 для бесконечного случая.

Утверждение 3. Если $AB = BA$, то $B^* \infty A^*$.

Доказательство. Предположим, что $A \not\ni e$. Рассматривая любое слово $v \in B^*$, выберем $l \in N_0$, для которого $v \in B^l$; обозначим $k = |v| + 1$. Таким образом, выполняется следующее условие:

$$v \in \text{opref}(B^l A^k).$$

Согласно доказанному выше, $B^l A^k = A^k B^l$, поэтому

$$v \in \text{opref}(A^k B^l).$$

Так как $e \notin A$, то мы имеем:

$$\|A^*\|_{\min} \geq k > |v| \quad \text{и} \quad v \in \text{opref}(A^k).$$

Рассмотрим произвольное слово $v \in B^*$. Так как $e \notin A$, то выполняется условие:

$$B^* \subseteq \text{opref}(A^*)$$

(т.е. $B^* \infty A^*$.)

Далее в этом утверждении предполагаем, что $A \ni e$.

Предположим, что $B^* \not\subseteq \text{opref}(A^*)$, т.е.

$$(\exists v \in B^*) (v \notin \text{opref}(A^*)).$$

Выберем $l \in N$ так, что $v \in B^l$. Рассмотрим также язык

$$D = \{w \in A^* | w \in \text{pref}(v)\}.$$

По способу построения D является конечным.

Заметим, что

$$v \notin \text{oref}(A^*),$$

тогда $v \notin D$, поэтому можно определить D другим способом, т.е.

$$D = \{w \in A^* | w \in \text{oref}(v)\}.$$

Так как D является конечным, то можно выбрать некоторое слово $u \in \max(D)$. Пусть выполняется следующее условие:

$$u \in A^{k_1} \quad \text{и} \quad |u| = m.$$

(Мы выбираем здесь k_1 и m .)

Так как $A \neq \{e\}$, для каждого $i < j$ выполняется неравенство:

$$A^i < A^j.$$

Поэтому можно выбрать некоторое $k_2 \in N$, такое, что

$$|A^{k_2}|(m+1) \cdot |B^l|.$$

Как и при доказательстве утверждения 2, обозначим

$$k = \max(k_1, k_2).$$

Заметим, что согласно условиям $e \in A$ и $u \in A^{k_1}$, слово u принадлежит языкам A^i для каждого $i \geq k_1$, поэтому $A^k \ni u$.

Обозначим

$$v = ua_1a_2 \dots a_n,$$

где

$$a_1, a_2, \dots, a_n \in \Sigma,$$

и, очевидно, $n > 0$. Определим конечные языки

$$C_1 = \{u, ua_1, ua_1a_2, \dots, ua_1a_2 \dots a_{n-1}\},$$

$$C_2 = v\Sigma^*, \quad C = C_1 \cup C_2.$$

Рассмотрим язык $D = A^k B^l$. Согласно разделу , равенство $D = B^l A^k$ тоже выполняется. Посчитаем двумя способами число различных элементов множества $D \cap C$.

С одной стороны, рассмотрим D как $D = A^k B^l$. Все слова uw , где $w \in B^l$, могут принадлежать $D \cap C$. Поэтому, согласно предположению $|u| = m$, слово u имеет ровно m различных собственных префиксов (включая e), то условие

$$u'w \in D \cap C$$

выполняется для любого слова u' последнего множества (т.е. $u' \in \text{oref}(u)$) и некоторое $w \in B^l$ выбирается по u' .

По определению отношения $\succ$ для языка $D \cap C$, получаем следующее условие:

$$|D \cap C| > |B^l| + m \cdot |B^l| = (m+1) \cdot |B^l|. \quad (6)$$

(Мы действительно посчитали все слова множества $D \cap C$, так как условия $u'' \in C$, $u'' \in A^k$ и $u'' \neq u$ не могут выполняться в силу следующего. Если u'' принадлежит языку C_1 , то мы получаем противоречие по способам выбора u как элемента множества $\text{max}(D)$, так как

$$u'' \in A^* \text{ and } |u''| \succ |u|.$$

И если условие $u'' \in C_2$ выполняется, то получается, что $v \in \text{oref}(A^*)$, что противоречит способу выбора v .)

Таким образом, не существует непересчитанных элементов языка $D \cap C$, т.е. неравенство (6) выполняется.

С другой стороны, рассмотрим язык D как $D = B^l A^k$. Вследствие того, что $A \in e$, условие $v \in B^l$ выполняется, поэтому все элементы множества D типа vw (где $w \in A^k$) принадлежат также C_2 , и, следовательно, C . Таким образом, число различных элементов языка $D \cap C$ не меньше, чем число различных элементов A^k . И, согласно условию $k \geq k_2$ и способу выбора k_2 , мы имеем

$$|D \cap C| > (m+1) \cdot |B^l|. \quad (7)$$

Итак, мы получили противоречие, (6) и (7), поэтому

$$B^* \subseteq \text{oref}(A^*),$$

т.е. $B^* \subseteq A^*$. □

Следующая теорема выполняется также и для бесконечного случая.

Теорема 2. Если $AB = BA$, то $A^* \approx B^*$. □

Случай префиксных языков

Далее в данной статье рассматриваются другой частный случай равенства (1), а именно, когда один из рассматриваемых языков содержит не более двух слов.

Рассмотрим равенство (1) в префиксном случае, т.е. пусть выполнены условия $\text{Pr}(A)$ и $\text{Pr}(B)$.

Теорема 3. Если $AB = BA$, $\text{Pr}(A)$ и $\text{Pr}(B)$, то

$$(\exists C \in \Sigma^*, k, l \in \mathbb{N}_0) (A = C^k, B = C^l).$$

Доказательство. Если $A = B$, то доказываемое утверждение верно при

$$C = A = B \text{ и } k = l = 1.$$

А если $A = \{e\}$ (в префиксном случае это равносильно тому, что $A \ni e$), то утверждение верно при

$$C = B, \quad k = 0, \quad l = 1;$$

аналогично для $B = \{e\}$.

Рассмотрев эти случаи, ниже в доказываемой теореме мы можем предполагать, что $A \neq B$, причём ни одно из этих множеств не содержит e (т.е. не равно $\{e\}$). При этом ограничении выполнено по крайней мере одно из двух следующих утверждений:

$$(\exists u \in A, v \in B) (u \in \text{opref}(v)), \quad (8)$$

$$(\exists u \in A, v \in B) (v \in \text{opref}(u)) \quad (9)$$

(мы пока не утверждаем, что выполнено *ровно* одно из них).

Пусть выполняется (8) (если (9), то рассуждения аналогичны). Для каких-либо выбранных согласно условию (8) слов u и v рассмотрим $w \in \Sigma^*$ такое, что $v = uw$. Выполнено условие

$$uwA \subseteq BA = AB,$$

поэтому $wA \subseteq B$.

Объединение всех слов w , которые могут быть выбраны по зафиксированному $u \in A$ и некоторому удовлетворяющему (8) слову $v \in B$, обозначим D . Согласно вышеизложенному, $DA \subseteq B$; поэтому в случае $DA \neq B$ условие (1) не выполняется вследствие $\text{Pr}(B)$. Итак,

$$D \cdot A = B, \quad (10)$$

причём условие (9) всё же не выполняется.

Согласно (1) и (10), выполнено равенство

$$ADA = DAA,$$

поэтому вследствие $\text{Pr}(A)$ верно следующее:

$$AD = DA \text{ и } \text{Pr}(D).$$

Таким образом, условия доказываемой теоремы выполняются и для языков A и D (вместо A и B), причём, поскольку $e \notin A \cup B$, верно также следующее:

$$\|A \cup D\|_{\min} < \|A \cup B\|_{\min} \quad (11)$$

(мы воспользовались простейшим неравенством, связывающим числа $\|A\|_{\min}$ в двух языках).

Обозначим теперь

$$A_0 = A_1 = A, \quad B_0 = B, \quad B_1 = D.$$

Если $A_1 \neq B_1$, то применив всё изложенное выше к языкам A_1 и B_1 , получим некоторые A_2 и B_2 и т.д. Согласно вводимым таким образом обозначениям языков A_i и B_i и обобщению утверждения (10) (перепишиваемому в общем случае – для вновь вводимых множеств – либо в виде $B_i A_i = B_i$, либо в виде $B_i A_i = B_i$), для всех возможных i выполнены неравенства

$$\|A_{i+1} \cup B_{i+1}\|_{\min} < \|A_i \cup B_i\|_{\min},$$

аналогичные (11). Из этих неравенств следует, что описанный процесс построения пар множеств A_i и B_i конечен, т.е. для некоторого $m \in \mathbb{N}$ верно следующее:

$$A_0 \neq B_0, \quad A_1 \neq B_1, \quad \dots, \quad A_{m-1} \neq B_{m-1}, \quad A_m = B_m.$$

Обозначим

$$C = A_m = B_m.$$

Согласно методу построения C ,

$$\{A_{m-1}, B_{m-1}\} = \{C, C^2\},$$

отсюда

$$\{A_{m-2}, B_{m-2}\} \subseteq \{C, C^2, C^3\},$$

и так далее. Таким образом, для некоторых натуральных k и l выполняются равенства

$$A = C^k, \quad B = C^l.$$

Ранее уже был рассмотрен случай, когда одно из множеств равно $\{e\}$; объединяя эти два случая, получаем, что для выбранного нами языка C ,

$$(\exists k, l \in \mathbb{N}_0) (A = C^k, B = C^l),$$

что и требовалось доказать. □

Замечание. В случае конечного алфавита Σ доказанная теорема имеет следующую числовую интерпретацию.

Во-первых, согласно теореме 1, можно считать, что множества A и B содержат некоторые максимальные префиксные коды (т.е.

$$A, B \in \text{pr}^+(\Sigma);$$

см. на эту тему также следующий подраздел). А поскольку мы рассматриваем случай префиксных множеств, то сами A и B являются максимальными префиксными кодами, т.е.

$$A, B \in \text{mp}(\Sigma).$$

Во-вторых, упорядочим алфавит Σ , т.е. рассмотрим какое-нибудь отношение строгого порядка " $\prec$ ": если

$$\Sigma = \{a_1, a_2, \dots, a_n\},$$

то положим, например,

$$a_1 \prec a_2 \prec \dots \prec a_n.$$

С помощью введённого отношения " $\prec$ " упорядочим и слова из Σ^* – т.е. будем рассматривать их в лексикографическом (алфавитном) порядке. Любые языки будем здесь рассматривать только в указанном порядке, т.е. пусть наши множества A и B суть

$$A = (u_1, u_2, \dots, u_k), \text{ где } u_1 \prec u_2 \prec \dots \prec u_k,$$

$$B = (v_1, v_2, \dots, v_l), \text{ где } v_1 \prec v_2 \prec \dots \prec v_l$$

(мы употребляем круглые скобки, поскольку рассматриваем A и B как упорядоченные множества, т.е. конечные последовательности). Соответствующие последовательности длин слов обозначим

$$\mathcal{D}(A) = (|u_1|, |u_2|, \dots, |u_k|), \quad \mathcal{D}(B) = (|v_1|, |v_2|, \dots, |v_l|).$$

Несложно убедиться, что поскольку порядок букв алфавита заранее зафиксирован, то произвольное множество $A \in \text{mp}(\Sigma)$ однозначным образом восстанавливается по построенной указанным способом последовательности длин $\mathcal{D}(A)$ (т.е. описанное отображение из множества максимальных префиксных кодов в множество числовых последовательностей является инъекцией).

По определению максимальных префиксных кодов несложно показать, что и

$$AB \in \text{mp}(\Sigma)$$

(т.е. множество максимальных префиксных кодов является подмоноидом рассматриваемого глобального надмоноида свободного моноида). Поэтому мы можем употреблять обозначение $\mathcal{D}(AB)$ в только что определённом смысле. Легко доказывается, что

$$\mathcal{D}(AB) = (|u_1| \cdot |v_1|, \dots, |u_1| \cdot |v_l|, \\ |u_2| \cdot |v_1|, \dots, |u_2| \cdot |v_l|, \dots, |u_k| \cdot |v_1|, \dots, |u_k| \cdot |v_l|).$$

А отсюда следует, что условие (1) равносильно равенству

$$\mathcal{D}(AB) = \mathcal{D}(BA).$$

Заключение

В данном разделе рассматриваются некоторые интересные примеры и задачи для дальнейшего решения.

Рассмотрим некоторые примеры. В префиксном случае все пары языков A и B , для которых выполнено условие (1), могут быть построены на основе теоремы 3. Очевидно, что (1) верно также для произвольных языков A и B , каждый из которых может быть записан в виде

$$\bigcup_{i \in I} \Sigma^i$$

для некоторого множества индексов $I \subseteq \mathbb{N}_0$.

Приведём менее очевидный пример непрефиксных A и B :

$$\Sigma = \{0, 1\},$$

$$A = \Sigma \cup \{00, 01, 11\} \cup \Sigma^3, \quad B = \Sigma \cup \Sigma^2; \quad (12)$$

тогда

$$AB = BA = \bigcup_{2 \leq i \leq 5} \Sigma^i.$$

Все примеры (эти и любые другие) могут быть изменены с помощью применения к языкам A и B одного и того же морфизма, причём не обязательно инъективного.

Далее сформулированы две гипотезы, для которых у авторов в настоящее время нет ни доказательств, ни контрпримеров.

Первая гипотеза состоит в следующем: если выполняется (1), то выполнено равенство

$$\text{pv}(A) \cdot \text{pv}(B) = \text{pv}(B) \cdot \text{pv}(A). \quad (13)$$

Заметим, что для произвольного языка

$$A \in \text{mp}^+(\Sigma),$$

согласно определению множества mp , выполнено условие

$$\text{pv}(A) \in \text{mp}(\Sigma).$$

Поэтому если первая гипотеза верна (т.е. если равенство (13) действительно выполняется для произвольных коммутирующих языков A и B), то, согласно теореме 1,

$$\text{pv}(A) = C^k \quad \text{и} \quad \text{pv}(B) = C^l$$

для некоторых $C \subseteq \Sigma^*$ (префиксного множества) и чисел $k, l \in \mathbb{N}_0$.

На основе этих равенств формулируется вторая гипотеза: при сделанных обозначениях можно выбрать множество C таким образом, что $A, B \subseteq C^*$.

(Заметим, что, например, пары бесконечных языков

$$A = \Sigma^2 \Sigma^* \quad \text{и} \quad B = \Sigma^4 \Sigma^*$$

для произвольного алфавита Σ контрпримерами ко второй гипотезе не являются, поскольку мы можем выбрать не только $C = \Sigma^2$, но и $C = \Sigma$.)

Для анализа обеих сформулированных гипотез может оказаться полезным следующий факт: вообще говоря,

$$\text{pv}(A) \cdot \text{pv}(B) \neq \text{pv}(A \cdot B). \quad (14)$$

Для подтверждения (14) рассмотрим такой пример:

$$\Sigma = \{0, 1\}, \quad A = B = \{0, 01, 100, 101, 110, 111\};$$

заметим, что здесь

$$A, B, AB \in \text{tr}^+(\Sigma).$$

Неравенство (14) при этом можно проверить непосредственно.⁴

Отметим в заключение, что с задачей описания критериев равенства $AB = BA$, а также со многими из задач, рассматривавшихся в предыдущих разделах и в [2, 4], непосредственно связана задача “извлечения корня” из заданного языка: для заданного языка $A \subseteq \Sigma^*$ требуется найти максимально возможное $n \in \mathbb{N}$ и зависящий от n язык $B \subseteq \Sigma^*$, такие что $A = B^n$. До конца эта задача авторами ещё не исследована, поэтому приведём лишь один интересный пример, имеющий некоторую аналогию с примером (12): для произвольного алфавита Σ (в том числе – при $|\Sigma| = 1$ и $|\Sigma| = \omega$) из языка

$$\bigcup_{2 \leq i \leq 10} \Sigma^i$$

извлекается не только “очевидный” квадратный корень

$$\bigcup_{1 \leq i \leq 5} \Sigma^i,$$

⁴ А можно показать и почти без “вычислений”, используя следующие соображения:

$$\text{pv}(A) = \text{pv}(B) = \{0, 100, 101, 110, 111\},$$

значит,

$$\text{pv}(A) \cdot \text{pv}(B) \not\subseteq 010,$$

в то время как $\text{pv}(AB) \ni 010$.

но и

$$\bigcup_{i \in \{1,2,4,5\}} \Sigma^i.$$

Таким образом, в случае языков операция извлечения корня заданной степени не является (однозначной) функцией.

Литература

- [1] А. Саломеа Жемчужины теории формальных языков, 1979.
- [2] B. Melnikov Some equivalence problems for free monoids and for subclasses of the CF-grammars class. – In: Number theoretic and algebraic methods in computer science, World Sci. Publ., 1995, pp. 125–137.
- [3] A.V. Aho, R. Sethi, J.D. Ullman Compilers: Principles, Techniques and Tools, N.J., Addison-Wesley, 1985.
- [4] B. Melnikov The equality condition for infinite catenations of two sets of finite words. – Int. J. of Found. of Comp. Sci. 4 3 (1993) pp. 267–274.

Методы единого поиска на примере игры «Сокобан».

М.В. Бузин

Данная статья является более подробным изложением результатов работы опубликованных в [1].

Исследование методов поиска является фундаментальным для Искусственного Интеллекта. Не вступая в рассуждения о том, что такое интеллект и как он может быть реализован, несомненным остается тот факт, что поисковые программы могут решить задачу, которая для своего решения, по мнению человека, требует интеллекта. Игры и головоломки являются примерами таких задач. Поэтому не случайно, что традиционно моделирование игр и головоломок является одним из основных направлений развития ИИ. По мнению некоторых психологов, человек сам формирует свой интеллект через игры. В первую очередь под пристальное изучение стали попадать игры и головоломки с конечным числом ситуаций и четко определенными правилами. Среди них можно выделить головоломки: «Волк, коза, капуста», «Пятнашки», «Кубик Рубика»; игры: «Го», «Нарды», «Шашки», «Шахматы». Здесь приведены известные игры и головоломки, которые очень сильно различаются по своей сложности, однако многие подходы к их решению одинаковы. Несомненным лидером по изученности и количеству публикаций являются шахматы¹. Тот факт, что сейчас уровень машины при игре в шахматы превосходит уровень человека, широко известен.

Целью данной работы является рассмотрение и анализ головоломки «Сокобан». Головоломка «Сокобан» несколько не уступает по сложности шахматам (хотя само сравнение игры и головоломки может быть только весьма условным). Простые правила, сочетаются с различной степенью сложности задач (от простых до сложных). Это делает Сокобан не только заманчивым предметом исследования в области моделирования игр, но и широко популярной во всем мире игрой. Помимо этого, Сокобан может иметь и некоторую практическую ценность: при наличии соответствующей робототехники разработанные и отлаженные на Сокобане методы позволяют полностью автоматизировать погрузочно-разгрузочные работы на складе.

Характеристика Сокобана.

К сфере искусственного интеллекта (ИИ) некоторые из ученых относят те различные области, где мы действуем, не имея абсолютно точного метода решения проблемы. Такие ситуации означают выбор в условиях неопределенности. Хотя понятие интеллекта до сих пор еще не определено и является предметом споров, можно рискнуть сказать, что именно этот недетерминизм, эта свобода действия и являются существенной составляющей интеллекта. Поэтому чем менее определен процесс поиска решения, тем большее значение имеет результат такого поиска.

Большинство описываемых в литературе исследований в области моделирования головоломок чаще всего затрагивают наиболее простые из существующих известных головоломок (например «Кубик Рубика», «Пятнашки»). Наиболее простые в том смысле, что они имеют некоторые или все из ниже перечисленных свойств.

1. Умеренная длина решения.
2. Незначительный фактор ветвистости дерева поиска.
3. Достаточно эффективная и недорогая оценка длины решения.
4. Отсутствие неразрешимых ситуаций.

Присутствие в характеристике задачи любого из этих свойств значительно упрощает процесс поиска решения. «Умеренная длина решения» и «Незначительный фактор ветвистости дерева поиска» позволяют эффективно использовать методы не информированного поиска. «Эффективная и недорогая оценка длины решения» позволяет свести весь процесс поиска к простому движению от исходного состояния к решению. «Отсутствие неразрешимых ситуаций» означает разрешимость из каждой позиции. Следовательно, нельзя ошибиться при выборе следующего шага, поэтому пусть и не оптимальное, но решение все равно будет получено. Сокобан не обладает ни одним вышеперечисленным свойством, и, более того, можно указать три основные характеристики, описывающие его сложность:

1. Огромное дерево поиска
2. Сложная оценка длины решения
3. Обилие тупиковых ситуаций

Опишем все эти свойства более подробно.

Огромное дерево поиска.

Говоря об огромных размерах дерева поиска решения в Сокобане, следует отметить три следующих фактора, определяющих объем и сложность производимого по дереву поиска. Во-первых, это огромная глубина дерева полного перебора. Вследствие относительной свободы при смене состояний поиск может использовать практически все доступное пространство. Размер этого пространства и количество размещенных на нем ящиков задают огромное количество вариантов выбора, которое не соизмеримо с возможностями полного перебора.

В качестве примера можно рассмотреть классическую задачу изображенную на рис. №1. Это первый лабиринт, предлагаемый Сокобаном для решения. Именно с него

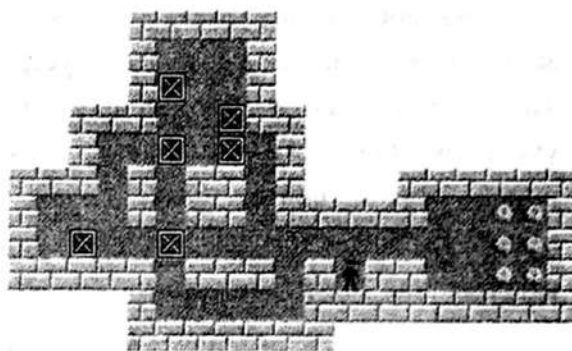


Рис. 1 Пример лабиринта

¹ Среди наиболее известных можно назвать: Г. М. Адельсон-Вельский. В. Л. Арлазаров, М. В. Донской: Программирование игр, М., Наука, 1978; Михаил Ботвинник: Алгоритмы игры в шахматы, М., Наука, 1968.

игрок начинает знакомиться с игрой. Стоит заметить, что даже эта задача становится непреодолимым препятствием для некоторых новичков.

Второй фактор, который существенно усиливает эффект первого - это значительный фактор ветвистости дерева поиска. Для произвольной задачи фактор ветвистости приблизительно равен $4N$, где N - количество ящиков. Исходя из того, что средний лабиринт Сокобана содержит от 6 до 12 ящиков, получаем средний фактор ветвистости приблизительно равным 32, в то время как для игры пятнашки он составляет не более 4, а для Кубика Рубика не более 18.

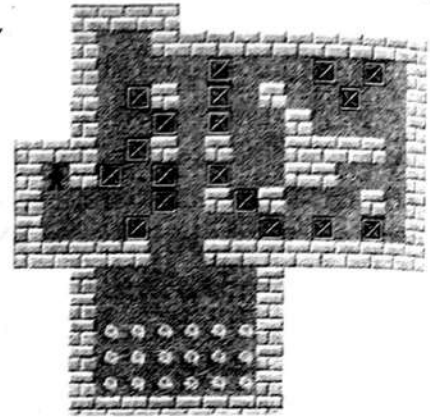


Рис 2. Пример длинного решения.

Помимо названных, существует третий фактор - огромное количество циклов. Хотя само заикливание может быть легко устранено путем запоминания уже рассмотренных состояний. Но, учитывая количество состояний, участвующих в поиске, такой подход может быть использован весьма условно, поскольку в случае достаточно глубоко расположенного решения даже современной ЭВМ будет сложно оперировать таким массивом данных.

Оценка длины решения.

В отличие от пятнашек, где решение не превосходит 80 шагов и Кубика Рубика, где длина решения ограничена 20 шагами (это доказано в [18]), длина решения в Сокобане доходит до 600 шагов (Например, решение лабиринта, изображенного на рис. №2, составляет 554 движения). Поэтому для Сокобана особенно важно, приступая к решению получить хотя бы приблизительную оценку ожидаемой длины будущего решения. Т.к. Сокобан имеет огромный фактор ветвистости, то даже один лишний шаг в глубину может привести к потере решения вообще.

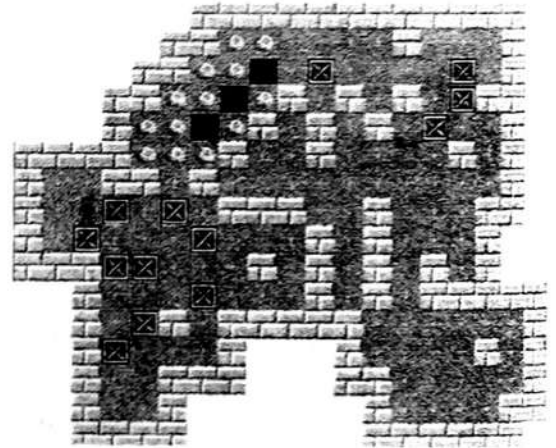


Рис. 3. Сокобан, лабиринт № 50.

Будет не верно полагать, будто бы вычисление длины решения сводится к суммированию всех расстояний от изначального расположения ящиков, до их предполагаемых мест хранения. Это не так, прежде всего потому, что, во время решения движение одного ящика по кратчайшему пути, как правило, затруднено другими ящиками (см. раздел «Конфликт продвижения»), либо в самой задаче заложена неразрешимость при передвижении кратчайшим путем. Например, на рис. №3 приведен пример задачи, решение которой предполагает перемещение большей части ящиков через места хранения.

В отличие от большого числа головоломок, для Сокобана не существует не только достаточно эффективной оценки длины решения, но и приемлемого алгоритма хотя бы для приблизительного ее вычисления. Т.о., само вычисление длины решения предстает в виде некоторого нетривиального поиска, который должен непосредственно предшествовать основному поиску решения, и от результата которого в значительной степени зависит успех основного решения.

Тупики.

Большинство рассматриваемых различными авторами головоломок имеют следующее свойство: начиная с разрешимого состояния каждое движение, приводит также к разрешимому состоянию. Это свойство согласно Корфу [19], гарантирует возможность нахождения некоторого (пусть даже и не оптимального) решения. Рассматривая с этой позиции Сокобан интересен тем, что не обладает этим свойством.

Процесс поиска решения в Сокобане постоянно заходит в тупиковые состояния. И только очень не многие тупики могут быть идентифицированы сразу. Тупиковая ситуация возникает всякий раз, когда хотя бы один ящик попадает в такое положение из которого он уже никогда не сможет достичь места хранения. Само наличие тупиков в процессе поиска создает неопределенность, которая лишает уверенности в правильности выбранного направления и ведет к коренному усложнению решения задачи.

Поводя итог вышесказанному, постановка задачи может быть сформулирована следующим образом. Огромное дерево доступных состояний. В этом дереве, постоянно избегая заикливания, необходимо отыскать достаточно глубоко расположенное решение. Причем весь процесс решения происходит в полной неопределенности о наличии впереди решения (неопределенность: поскольку практически невозможно определить насколько текущее состояние близко к решению и содержит ли текущее поддерево решение или нет).

Классические методы решения задач

Для поиска решения может быть предложено много различных алгоритмов. В данной статье будут рассмотрены только некоторые из них. Поскольку в качестве модели исследования выбрана головоломка Сокобан, поэтому дальше будут затронуты только методы «единого поиска»². Единый поиск предполагающий существование только одного процесса стремящегося к единой цели, используется в задачах без внешнего влияния, таких как: планирование; игры для одного человека; головоломки. Далее будут рассмотрены две группы методов относящихся к единому поиску: не информированные и информированные.

² Помимо единого поиска еще существует противоборствующий поиск. Противоборствующий поиск состоит из двух или более процессов поочередно пытающихся достичь противоположных целей. Этот поиск используется при решении игр для двух или более человек. Например, метод минимакса при игре в шашки и шахматы

Не информированный поиск.

Случайный поиск.

Процесс поиска состоит из случайного выбора следующего шага. Метод легко реализуем и не требует системного подхода к анализу задачи. Этот метод может быть вполне результативным при высокой степени плотности решения.

Поиск в ширину.

Процесс поиска сначала исследует все вершины ближайшие к стартовому состоянию, только потом переходя к более дальним вершинам. Медленное углубление поиска и его быстрое расширение приводят к быстрому истощению системных ресурсов, что в свою очередь ограничивает применимость этого метода. В тоже время этот метод легок в реализации и обеспечивает оптимальное решение.

Поиск в глубину.

Процесс поиска следует вдоль дерева поиска сверху вниз до самой последней вершины, только потом исследуются все оставшиеся нерассмотренными поддеревья. Вообще говоря, поиск в глубину не обеспечивает оптимального решения, но по сравнению с поиском в ширину имеет неоспоримое преимущество – требуемое для поиска пространство. Поиск в глубину сохраняет только соседние вершины и поэтому требуемое для поиска пространство возрастает линейно. Поиск в ширину хранит все вершины вплоть до исследуемого уровня, а, следовательно, требуемое для поиска пространство возрастает экспоненциально.

Постепенно углубляющийся поиск.

Этот поиск возник, как попытка исправить некоторые несовершенства поиска в глубину, такие как цикл. Впервые этот метод поиска был представлен в [17, 23]. Основная идея заключается в постепенном увеличении максимальной глубины поиска. Если решение не найдено, то максимальная глубина увеличивается и поиск начинается заново. В отличие от поиска в глубину, постепенно углубляющийся поиск обеспечивает оптимальное решение.

Постепенно расширяющийся поиск.

Этот метод, базирующийся на поиске в ширину, был предложен Гинзбергом в [8]. Идея поиска заключается в постепенном увеличении числа рассматриваемых в каждом узле поддеревьев. Если поиск оказывается неудачным, то число рассматриваемых поддеревьев увеличивается и поиск начинается заново. Важно заметить, что хотя рост пространства поиска значительно уменьшился, но все же остался экспоненциальным. Помимо этого по сравнению с поиском в ширину, постепенно расширяющийся поиск теряет оптимальность решения и приобретает циклические процессы.

Поиск лучом.

Еще одно усовершенствование поиска в ширину – поиск лучом было предложено [24]. В отличие от постепенно расширяющегося поиска, идея поиска лучом состоит в

ограничении числа рассматриваемых поддеревьев не в каждом узле, а на всем уровне. Такой подход обеспечивает линейный прирост пространства поиска, но эффективность его использования весьма сомнительна.

Информированный поиск.

Перед тем как приступить к рассмотрению методов не информированного поиска, необходимо сказать несколько слов об используемой в дальнейшем терминологии. Точное расстояние от текущего состояния s до ближайшей цели будет определяться функцией $h^*(s)$. Поскольку, как правило, $h^*(s)$ неизвестна, то вводится некоторая оценка $h^*(s)$ – эвристическая функция $h(s)$. Причем $h(s)$ является допустимой, если $0 \leq h(s) \leq h^*(s)$, $\forall s \in S$, где S – множество всех состояний пространства поиска. $h(s)$ является последовательной, если $h(s_1) \leq h(s_2) + c(s_1, s_2)$, $\forall s_1, s_2 \in S$, $c(s_1, s_2)$ – стоимость перехода из s_1 в s_2 . Введем меру пройденного пути функцию $g(s)$. $g(s)$ это стоимость пути пройденного от исходного состояния s_0 до состояния s . Определим функцию $f(s)$ следующим образом: $f(s) = g(s) + h(s)$. Очевидно, что $f(s)$ является оценкой всего пути из s_0 в итоговое состояние s_g через состояние s . $f^*(s)$ определенное как $f^*(s) = g(s) + h^*(s)$, является действительной стоимостью перемещения по оптимальному пути из s_0 в s_g через состояние s .

Простой выбор лучшего.

Этот поиск основан на простом выборе следующей лучшей вершины. «Лучшая» вершина определяется как некоторая мера, обычно оценивается расстояние до ближайшей цели $h(s)$. К сожалению, этот метод не обеспечивает оптимального решения, поскольку не учитывает стоимость пройденного пути.

A*.

A* был впервые описан в [9]. Это первый метод, который использует $f(s)$ в качестве меры «лучшего». Учитывая стоимость пройденного пути и оценку оставшейся части пути, A* гарантирует получение оптимального решения. Интересно, что при таком подходе автоматически исключается заикливание.

IDA*.

Корф в [17] применяет идею постепенного углубления к методу A*. Основная идея заключается в ограничении глубины поиска некоторой величиной. Для первой итерации максимальная глубина поиска устанавливается равной $h(s)$. Если поиск оказывается неудачным, то глубина отодвигается и поиск повторяется заново. Ограничение длины решения, позволяет обеспечить дополнительный контроль величины дерева поиска.

Метод ветвей и границ в глубину.

Данный метод³ поверхностно очень напоминает только что рассмотренный IDA*. Также как и в IDA* в методе ветвей и границ в глубину (МВГГ) глубина решения огра-

³ Описан в [20]

ничивается некоторой величиной, и выбор следующего определяется функцией $f(s)=g(s)+h(s)$. Но на самом деле эти методы принципиально различны. В отличие от IDA*, в MBIT для начальной глубины решения выбирается уже не нижняя граница решения, а верхняя. Т.о. процесс поиска не делится на итерации, а граница решения постепенно поднимается до оптимального решения, через промежуточные решения (это означает, что при получении решения поиск не останавливается, а только поднимает границу решения и идет дальше). Другим отличием является порядок выбора следующего. Хотя и IDA* и MBIT при выборе руководствуются значением функции $f(s)$, но в отличие от IDA*, рассчитывающего $f(s)$ для всех открытых состояний, MBIT аналогично поиску в глубину, выбирает только среди открытых состояний текущей вершины.

Двунаправленный поиск.

Почему процесс поиска должен идти к решению, а не от него? И почему надо искать только в одном направлении? Действительно, интуитивно понятно, что одновременный поиск с двух сторон будет значительно результативней обычного одностороннего поиска. Двунаправленный поиск [21] создает два противоположных дерева поиска, в процессе углубления поиска ищется пересечение этих деревьев, которое и даст решение. Несмотря на очевидную перспективность, этот подход до сих пор еще мало изучен и почти не используется. Возможно, это связано с трудностями реализации такого подхода. Достаточно сложно определить место и сам факт пересечения двух процессов поиска. При реализации обратного поиска необходимо учитывать, что конечное состояние не обязательно единственно. И, наконец, само кодирование параллельного поиска требует дополнительных навыков.

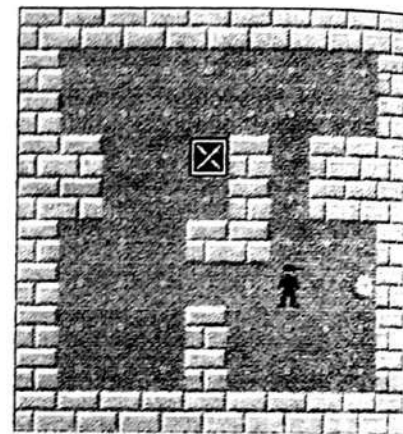


Рис. 4.

Какой метод выбрать?

Выбор того или иного метода в основном зависит от конкретной задачи, решение которой необходимо получить. Как правило, решающими критериями являются: ветвистость, глубина дерева поиска, и имеющиеся знания о предмете решения. Идеально было бы иметь метод, не требующий присутствия в нем предметно зависимых формул. В случае получения такого метода стало бы возможным решение произвольной задачи без предварительной и детальной ее проработки, без выделения специальных правил при ее решении. Часто встречающиеся в литературе описания различного рода «Универсальных решателей задач» являются результатом стремления к созданию такого метода.

Хотя универсальный метод достаточно перспективен, но на данном этапе существующие «универсальные решатели задач» эффективно работают только в достаточно простых проблемных областях, и не в состоянии оспаривать более сложные. Хорошим примером этого может служить испытание, описанное в [11]. Для испытания была выбрана программа победитель AIPS'984 «BlackBox» и два очень простых лабиринта головоломки Сокобан (рис. №4, рис.№5). Результат говорит сам за себя: программа, ориентированная на решение лабиринтов Сокобана, справилась с обеими задачами в течение секунды, в то время как программе «BlackBox» потребовалось несколько секунд на решение первого лабиринта и более часа на решение второго. Этот пример, очевидно, свидетельствует о том, что для достижения успеха при решении конкретной задачи необходим ее тщательный анализ и специальные ориентированные на нее методы.

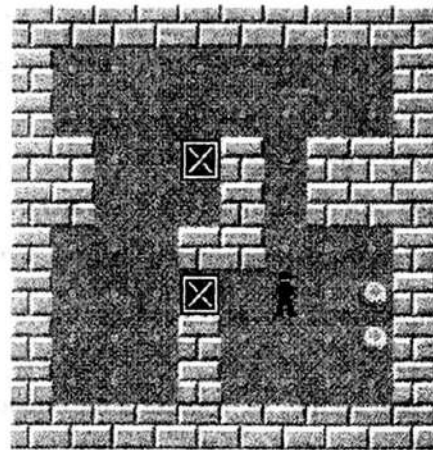
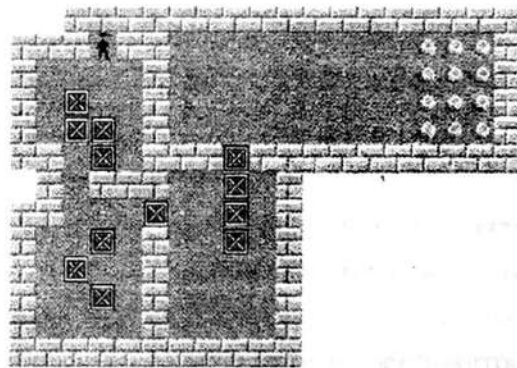


Рис. 5.

Специальные методы решения

Приступая к анализу возможного решения головоломки Сокобан, важно сразу оговорить, что поставленной целью является получение любого (не обязательно оптимального) решения. Учитывая сложность задачи, приступать к поиску оптимального решения имеет смысл, только тогда, когда будут тщательно изучены и отработаны методы нахождения решения в принципе. В качестве базового метода для реализации решения был выбран метод ветвей и границ в глубину. Поэтому следующие предлагаемые методы, были специально разработаны, чтобы усовершенствовать и адаптировать МВГГ под решение задач Сокобана. Многие из них уже были апробированы на практике.



Приступая к решению методом ветвей и границ в глубину на первом этапе, вследствие огромных размеров дерева поиска, очень важно получение как можно более точной оценки длины решения.

Оценка длины решения.

Сущность задачи вычисления длины решения, может быть представлена следующим образом. Есть n ящиков и n целей. Для каждого ящика минимальное количество движений до каждой цели в принципе вычислимо. Следовательно, стоимость пере-

⁴ Artificial Intelligence Planning Systems

движения ящика на место, для каждой пары ящик, место определена. Задача в том, чтобы выделить такую последовательность пар ящик, цель, которая дает минимальную стоимость. Казалось бы, исходя из поставленных условий, задача может быть решена с помощью полного сравнения в двухчастичном графе, сложность которого равна $O(k^3 \log_{2+k/4} k)$. Но проблема остается в том, что требуется найти не просто набор пар, дающий минимальную стоимость, а последовательность таких пар. Необходим поиск именно последовательности. Кроме того, необходимо учитывать существование следующих ситуаций:

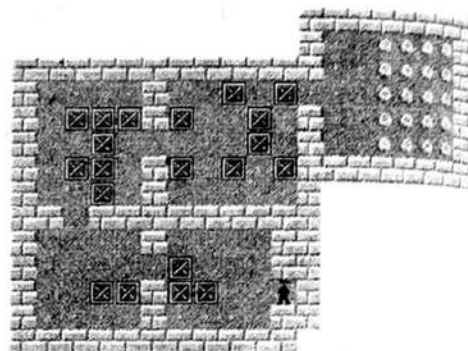


Рис. 7. Сокобан, лабиринт №4

Конфликт продвижения.

Рассмотрим ситуацию, когда два ящика препятствуют движению друг друга по оптимальному пути. Пример такой ситуации можно увидеть при решении лабиринта №80 (рис № 6). Очевидно, что для достижения цели, один из ящиков должен быть сдвинут с оптимального пути. Т.о. в случае если удастся идентифицировать такую ситуацию, то нижняя граница должна быть увеличена.

Петля.

Другой случай, когда так же требуется корректировка полученной стоимости длины решения – это ситуация изображенная на рис. №7. Поскольку докер только толкает ящики, и не может их тянуть, то для движения в определенном направлении докеру необходимо встать с соответствующей стороны от ящика. Часто встречаются такие ситуации, когда для достижения этой цели необходимо толкать ящик в противоположную сторону.

Сквозь цель

Это случай когда, для того чтобы все ящики могли быть расставлены по своим местам, необходимо один или более ящиков передвигать через целевую область, не задерживаясь в ней. Данный случай наиболее труден для моделирования, поскольку критерий достижения целевой области, является практически единственной оценкой близости решения.

Сложность вычисления оценки длины решения и относительная точность самого вычисления, не позволяют быстро и уверенно оперировать функцией $h(s)$ для выбора следующего шага. Поэтому для этих целей могут быть предложены следующие методы: инерционный и поиск по аналогии.

Инерционный метод.

Исходя из практических результатов, можно утверждать, что при введении описанного далее порядка выбора, процесс поиска приобретает разумные очертания, лишаясь

ется хаотичности и в значительной мере сокращает время, затрачиваемое на решение лабиринта. Предлагается следующий порядок выбора:

1. Выбираются движения, перемещающие ящики с одной целевой клетки на другую. Т.о. ящики уже достигшие целевой области, компонуются в ее пределах, освобождая место для других ящиков.
2. Выбираются движения, перемещающие на целевые клетки ящики еще них не находящиеся.
3. Обрабатываются все движения ящиков еще не достигших целевых клеток.
4. Выбираются движения, перемещающие ящики, находящиеся на целевых клетках в сторону от них.

Внутри каждого выбора наибольшим приоритетом обладают ящик, который двигался последним и направление, в котором осуществлялось движение этого ящика.

Данный метод вполне применим для большинства лабиринтов. Но, тем не менее, имеются такие лабиринты (например, лабиринт №50, рис. № 33), для которых данный метод потенциально отодвигает решение.

Поиск по аналогии

После достижения целевой клетки одним из ящиков можно попробовать использовать путь, пройденный этим ящиком, для передвижения оставшихся ящиков. Данный метод дает эффект только в случае лабиринтов, где целевые точки сгруппированы в одной комнате и в решении часто используются длинные передвижения.

Несмотря на рассчитанную оценку решения и произведенное на ее основании отсечение дерева поиска, оно все еще остается достаточно большим, чтобы позволить потеряться решению. Поэтому для дерева поиска необходимы дополнительные методы контроля и отсечения. Среди используемых методов контроля рассмотрим: Анализ повторения и анализ тупиков. А методы отсечения будут представлены методами: однозначный путь и уместность отсечения.

Анализ повторения

Рассматривая подробно дерево возможных состояний легко заметить, что при такой скорости разрастания вширь, дерево быстро становится плотно наполненным одинаковыми состояниями. Само присутствие в дереве перебора одинаковых состояний может привести к заикливанию. Чтобы избежать заикливания и попутно значительно сократить дерево перебора, можно использовать достаточно простой и надежный способ – запоминание разобранных состояний.

Используя метод запоминания, важно помнить, что сравнивая два состояния необходимо учитывать, не только положение ящиков, но и докера. Иначе говоря, два состояния являются идентичными (подобными) если положение ящиков совпадают, и существует путь из положения докера в первом состоянии в положение докера из второго состояния по которому, не двигая ящиков, может пройти докер.

К сожалению, несмотря на свою простоту и надежность данный метод имеет один существенный недостаток – ограниченность применения. Ограниченность применения возникает вследствие ограниченных ресурсов компьютера.

Анализ тупиков

Ситуация, когда имеется один или более ящиков, которые уже никогда не достигнут цели, или когда имеется цель, на которую уже никогда нельзя будет поставить ящик – выглядит явно как тупиковая. Описанная ситуация встречается довольно часто. К сожалению довольно сложно вовремя идентифицировать тупиковую ситуацию. Возможны следующие методы определения тупиковых ситуаций:

Комбинация базовых тупиковых ситуаций.

Выделяются очевидные базовые конфигурации тупиков. В процессе поиска для определения тупиковой ситуации в текущем состоянии ищутся базовые тупики или их комбинация.

Таблица тупиков.

Определяется некоторая область, внутри которой перечисляются все всевозможные конфигурации тупиков. В процессе поиска для определения тупиковой ситуации в текущем состоянии ищется соответствующих размеров и соответствующим содержанием область.

Поиск по образцу.

Суть этого метода заключается в периодической проверке на тупиковую ситуацию. Сама проверка происходит в виде параллельного поиска, суть которого подробно описана в [13].

Однозначный путь

Рассмотрим движение ящика через коридор. Слева стена, справа стена, остается либо вперед, либо назад (если докер сможет обойти коридор с другой стороны). Если внимательно посмотреть на это движение, то налицо ситуация, когда вне зависимости от расположения ящика внутри туннеля, ситуация не меняется. Т.о. весь туннель может быть представлен одной точкой. Туннель лишь частный случай общего варианта однозначно определенного пути. Данный метод значительно сокращает размер рассматриваемого пространства за счет преобразования вершины некоторых клик в цепочку.

Уместность отсечения

Рассматривая процесс поиска решения компьютером, сразу бросаются в глаза движения, которые человек бы никогда не допустил. Компьютер рассматривает даже совершенно не относящиеся друг к другу движения, только для того чтоб доказать, что данное поддерево не содержит решения. Предлагаемый метод, делает попытку внести разумность в процесс поиска. Суть метода заключается в ведении меры влияния, рассчитываемой для каждого ящика по отношению друг к другу, и уместности движения. Если два ящика не влияют друг на друга, то маловероятно, что они уместны, друг относительно друга. Данный метод, анализируя ситуацию, отбрасывает не уместные движе-

ния, оставляя только уместные. Более подробно об этом методе можно прочитать в [16].

Заключение

Из всего выше сказанного видно, что Сокобан действительно перспективная область исследования, и хорошая тестовая база для отработки методов ИИ на практике. Описанные выше методы решения – далеко не полный список возможных способов решения головоломки Сокобан. И, скорее всего после практического внедрения всех этих методов, они смогут быть доработаны дать впечатляющий результат.

Литература

- [1] Бузин М.В. *Сокобан как область применения методов ИИ* // Программирование (РАН), 1999, представлена к публикации.
- [2] Мельников Б.Ф., Бузин М.В. *К анализу игры Сокобан* // Труды второй международной конференции «Математическое моделирование физических, экономических, социальных систем и процессов», с. 39-40, 1999.
- [3] Бузин М.В., Яковлев В.И. *Математическая модель специальной области применения интеллектуального автомата* // Ульяновский филиал БАТТ, научно методический сборник №10, 1999.
- [4] Бузин М.В., Кузнецов В.А. *Анализ и методы решения в специальной области применения интеллектуального автомата* // Ульяновский филиал БАТТ, научно методический сборник №10, 1999.
- [5] Culberson J. *Sokoban is PSPACE-complete* 1997, на <http://www.cs.ualberta.ca/~joe/publications.html>
- [6] Dor D., Zwick U. *Sokoban and other motion planing problems* 1995, на <http://www.math.tau.ac.il/~ddorit/>
- [7] Finkelstein L., Markovitch S. *A selective macro-learning algorithm and its application to the N*N Sliding tile puzzle* Artificial Intelligence, 8, pp. 223-263, 1998
- [8] Ginsberg M. *Dynamic backtracking* // Journal of Artificial Intelligence Research, 1:25-46, 1993.
- [9] Hart P.E., Nilson N.J., Raphael B. *A formal basis for heuristic determination of minimum cost paths* // IEEE Transactions on System Science and Cybernetics, 4(2):100-107, 1968.
- [10] Holzinger W. *Sokoban – facing the infeasible*, 1998.
- [11] Junghanns A., Schaeffer J. *Domain-Dependent Single-Agent Search Enhancements* 1999, на <http://www.cs.ualberta.ca/~andreas/Papers/publications.html>
- [12] Junghanns A., Schaeffer J. *Relevance cuts: Localizing the search* 1998, на <http://www.cs.ualberta.ca/~andreas/Papers/publications.html>
- [13] Junghanns A., Schaeffer J. *Single-agent search in the presence of deadlocks* 1998, на <http://www.cs.ualberta.ca/~andreas/Papers/publications.html>

- [14] Junghanns A., Schaeffer J. Sokoban: *A challenging single-agent search problem* 1997, на <http://www.cs.ualberta.ca/~andreas/Papers/publications.html>
- [15] Junghanns A., Schaeffer J. Sokoban: *Evaluating standard single-agent search techniques in the presence of deadlock* 1998, на <http://www.cs.ualberta.ca/~andreas/Papers/publications.html>
- [16] Junghanns A., Schaeffer J. Sokoban: *Improving the Search with Relevance Cuts* 1999, на <http://www.cs.ualberta.ca/~andreas/Papers/publications.html>
- [17] Korf R. E. *Depth-first iterative-deepening: An optimal admissible tree search* // Artificial Intelligence, 27(1):97-109, 1985.
- [18] Korf R. E. *Finding optimal solutions to Rubic's Cube using pattern databases* // AAAI National Conference, pp.700-705, 1997.
- [19] Korf R. E. *Real-Time heuristic search* // Artificial Intelligence, 42(2-3), pp.189-211, 1990.
- [20] Lawler E.L., Woods D. *Branch-and-bound methods: A survey* // Operations Research, 14, 1966.
- [21] Pohl I. *Bi-directional Search* pp. 127-140. Edinburgh University Press, Edinburgh, 1971.
- [22] Murase Y., Matsubara H. and Hiraga Y. *Automatic Making of Sokoban Problems*, 1996
- [23] Slate D., Atkin L. *Chess 4.5 – The Northwestern University chess program* // Chess Skill in Man and Machine, pp. 82-118, New York, 1977.
- [24] Winston P.H. *Artificial Intelligence*, Addison-Wesley, 1992.
- [25] Лорьер Ж.-Л. *Системы искусственного интеллекта*, М. 1991

КРИТЕРИЙ КОНЕЧНОСТИ КОДЛИНЫ МНОГООБРАЗИЙ Z_2 -ГРАДУИРОВАННЫХ АССОЦИАТИВНЫХ АЛГЕБР

И.Р. Васильева

Данная статья посвящена многообразиям Z_2 -градуированных ассоциативных алгебр конечной кодлина. Центральным результатом статьи является критерий конечности кодлина для таких многообразий, выраженный в виде теоремы 1.

Введем некоторые понятия и обозначения.

На протяжении всей работы через K будем обозначать основное поле, нулевой характеристики. В случае поля нулевой характеристики [1] вся информация о многообразии, содержится в его полилинейной части, которая образует линейной подпространство в относительно-свободной алгебре многообразия. Полилинейная часть $P_n^{gr}(V)$ многообразия V супералгебр есть пространство полилинейных элементов степени n от n переменных:

$$P_n^{gr}(V) = \langle \{x_{i_1}^{g_{i_1}} \dots, x_{i_n}^{g_{i_n}} | g_{i_1}, \dots, g_{i_n} \in \{1, \phi\}\} \rangle,$$

где ϕ - автоморфизм порядка 2 относительно-свободной алгебры данного многообразия. Заметим, что $x^{(g_1 g_2)} = (x^{g_2})^{g_1}$, поскольку по определению $x^g = g(x)$.

Размерность $P_n^{gr}(V)$ называется n -й градуированной коразмерностью $c_n^{gr}(V)$ многообразия V и ее рост при $n \rightarrow \infty$ определяет рост самого многообразия.

Говорят, что рост многообразия V полиномиальный, если существуют такие константы s и r , что при любом n будет выполнено неравенство $c_n^{gr}(V) \leq sn^r$. Рост многообразия почти полиномиальный, если рост любого подмногообразия данного многообразия полиномиальный, в то время, как рост самого многообразия полиномиальным не является. Напомним, что в случае основного поля нулевой характеристики, многообразий роста между полиномиальным и показательным не бывает ([2]).

Полилинейную часть любого многообразия можно представить как левый модуль над некоторой конечной группой, вернее соответствующей групповой алгеброй. Так, пространство $P_n^{gr}(V)$ является H_n - модулем, где H_n - гипероктаэдральная группа порядка n , состоящая из элементов вида $h = (g_1, \dots, g_n; \sigma)$, $g_i \in G, i = 1, \dots, n$, $\sigma \in S_n$, мультипликативная группа $G = \{\phi, 1\}$ имеет порядок 2. Элементы группы H_n действуют на элементы пространства $P_n^{gr}(V)$ следующим образом (см [5]):

$$h(x_1 \dots x_n) = x_{\sigma(1)}^{g_{\sigma(1)}^{-1}} \dots x_{\sigma(n)}^{g_{\sigma(n)}^{-1}}.$$

Из теоремы Машке следует, что полилинейная часть $P_n^{gr}(V)$ многообразия супералгебр раскладывается в прямую сумму неприводимых подмодулей. Число составляемых в данном разложении назовем кодлыной многообразия. Поясним понятие кодлыны, используя язык кратностей и характеров.

Изучение пространства $P_n^{gr}(V)$ удобнее свести к изучению пространств $P_{r,n-r}^{gr}(V)$ от r четных переменных $y_i, y_i^\phi = y_i, i = 1, \dots, r$ и $n-r$ нечетных переменных $z_j, z_j^\phi = -z_j, j = 1, \dots, n-r$. Причем на четных переменных действует стандартный образ группы S_r , а на нечетных - группа S_{n-r} , а группа H_n действует на элементах пространства $P_{r,n-r}^{gr}(V)$ следующим образом:

$$hy_i = y_{\sigma(i)}, \quad hz_j = z_{\sigma(j)}^{\sigma(j)} = \pm z_{\sigma(j)}, \quad h = (g_1, \dots, g_n; \sigma) \in H_n.$$

Тогда, для многообразия V супералгебр, как в случае многообразий ассоциативных алгебр с инволюцией, n -й градуированный H_n -характер $\chi_n^{gr}(V)$ равен ([3] Theorem 1.3):

$$\chi_n^{gr}(V) = \sum_{r=0}^n \sum_{\substack{\lambda \vdash r \\ \mu \vdash n-r}} m_{\lambda, \mu} \chi_\lambda \otimes \chi_\mu,$$

где χ_λ (соответственно χ_μ) - неразложимый S_r -характер (соответственно S_{n-r} -характер), соответствующий разбиению λ (соответственно μ), $m_{\lambda, \mu} \geq 0$ - некоторые кратности.

И кодлыну многообразия V супералгебр можно определить как

$$l_n^{gr} = l_n^{gr}(V) = \sum_{r=0}^n \sum_{\substack{\lambda \vdash r \\ \mu \vdash n-r}} m_{\lambda, \mu}.$$

Если ввести обозначение $c_{r,n-r}^{gr}(V) = \dim P_{r,n-r}^{gr}(V)$, тогда n -ю градуированную коразмерность многообразия V супералгебр можно определить как ([6]):

$$c_n^{gr}(V) = \sum_{r=0}^n \binom{n}{r} c_{r,n-r}^{gr}(V).$$

Мищенко С.П., Джамбруно А. и Зайцевым М.В. ([6]) доказано, что существует только пять многообразий супералгебр почти полиномиального роста (Следствие 3. [6]). Дадим их краткое описание.

Многообразия, порожденные бесконечномерной Z_2 -градуированной алгеброй Грассмана G над полем K и Z_2 -градуированной алгеброй UT_2 верхнетреугольных матриц размера 2×2 над полем K с естественным типом градуировки $G = G \oplus 0$, $UT_2 = UT_2 \oplus 0$, будем обозначать $supvar G$ и $supvar UT_2$ соответственно. Строение полилинейных частей данных многообразий полностью совпадает со строением полилинейных частей многообразий $var G$ и $var UT_2$ ассоциативных алгебр (см [7], [8]). Отсюда следует, что их кодлына растет полиномиально.

Многообразия, порожденные алгебрами G и UT_2 с заданой на них стандартной градуировкой $G = G_0 \oplus G_1$, $UT_2 = Ke_{11} + Ke_{22} \oplus Ke_{12}$, где G_0 - линейное пространство мономов от e_i четной длины, а G_1 - нечетной длины. Мы будем обозначать их соответственно $supvarG^{gr}$ и $supvarUT_2^{gr}$.

Мищенко С.П., Джамбруно А. и Зайцевым М.В. ([6]) описан идеал тождеств многообразия $supvarG^{gr}$, полилинейная часть и приводятся формула и точная оценка n -й градуированной ко размерности (Предложение 3). Откуда следует, что кодлина также растет полиномиально $l_n^{gr}(supvarG^{gr}) = \sum_{r=0}^n 1 = n + 1$.

А. Валенти описала многообразие $supvarUT_2^{gr}$ и показала, что оно имеет полиномиальную кодлину.

И, наконец, многообразие супералгебр, порожденное алгеброй $K \oplus tK$, $t^2 = 1$. Благодаря коммутативности умножения на элементах алгебры G_2 , автоморфизм порядка 2 можно всегда рассматривать как инволюцию. Таким образом, строение полилинейной части данного многообразия супералгебр полностью совпадает со строением полилинейной части многообразия $varG_2$ ассоциативных алгебр с инволюцией [4]. Данное многообразие также имеет полиномиальный рост кодлины $l_n(varG_2, *) = n + 1$.

Из всего вышесказанного следует Замечание, которое понадобится нам при доказательстве критерия конечности кодлины многообразия супералгебр:

Замечание. Кодлины многообразий $supvarG$, $supvarUT_2$, $supvarG^{gr}$, $supvarUT_2^{gr}$ и $supvar(K \oplus tK)$, $t^2 = 1$ имеют полиномиальный рост.

Ниже мы сформулируем и докажем критерий ограниченности кодлины для многообразий ассоциативных Z_2 -градуированных алгебр константой. Для доказательства критерия нам понадобится следующая Лемма.

Лемма. Пусть многообразие V ассоциативных Z_2 -градуированных алгебр удовлетворяет условию $UT_2 \notin V$, тогда в нем выполняется тождество

$$y^T x y^{N-1-T} \equiv \sum_{i < T} \gamma_i y^i x y^{N-1-i} \quad (1)$$

где T, N - некоторые константы, коэффициенты $\gamma_i \in K$, переменная y - четная, а переменная x - произвольная.

Критерий конечности кодлины многообразий супералгебр гласит:

Теорема. Пусть V - многообразие ассоциативных Z_2 -градуированных алгебр над

полем нулевой характеристики. Тогда следующие условия эквивалентны:

- (i) многообразие имеет конечную кодлину;
- (ii) многообразие имеет полиномиальный рост.

Доказательство.

Многообразия, порожденные алгебрами G и UT_2 с заданой на них стандартной градуировкой $G = G_0 \oplus G_1$, $UT_2 = Ke_{11} + Ke_{22} \oplus Ke_{12}$, где G_0 - линейное пространство мономов от e_i четной длины, а G_1 - нечетной длины. Мы будем обозначать их соответственно $supvarG^{gr}$ и $supvarUT_2^{gr}$.

Мищенко С.П., Джамбруно А. и Зайцевым М.В. ([6]) описан идеал тождеств многообразия $supvarG^{gr}$, полилинейная часть и приводятся формула и точная оценка n -й градуированной ко размерности (Предложение 3). Откуда следует, что кодлина также растет полиномиально $l_n^{gr}(supvarG^{gr}) = \sum_{r=0}^n 1 = n + 1$.

А. Валенти описала многообразие $supvarUT_2^{gr}$ и показала, что оно имеет полиномиальную кодлину.

И, наконец, многообразие супералгебр, порожденное алгеброй $K \oplus tK$, $t^2 = 1$. Благодаря коммутативности умножения на элементах алгебры G_2 , автоморфизм порядка 2 можно всегда рассматривать как инволюцию. Таким образом, строение полилинейной части данного многообразия супералгебр полностью совпадает со строением полилинейной части многообразия $varG_2$ ассоциативных алгебр с инволюцией [4]. Данное многообразие также имеет полиномиальный рост кодлины $l_n(varG_2, *) = n + 1$.

Из всего вышесказанного следует Замечание, которое понадобится нам при доказательстве критерия конечности кодлины многообразия супералгебр:

Замечание. Кодлины многообразий $supvarG$, $supvarUT_2$, $supvarG^{gr}$, $supvarUT_2^{gr}$ и $supvar(K \oplus tK)$, $t^2 = 1$ имеют полиномиальный рост.

Ниже мы сформулируем и докажем критерий ограниченности кодлины для многообразий ассоциативных Z_2 -градуированных алгебр константой. Для доказательства критерия нам понадобится следующая Лемма.

Лемма. Пусть многообразие V ассоциативных Z_2 -градуированных алгебр удовлетворяет условию $UT_2 \notin V$, тогда в нем выполняется тождество

$$y^T x y^{N-1-T} \equiv \sum_{i < T} \gamma_i y^i x y^{N-1-i} \quad (1)$$

где T, N - некоторые константы, коэффициенты $\gamma_i \in K$, переменная y - четная, а переменная x - произвольная.

Критерий конечности кодлины многообразий супералгебр гласит:

Теорема. Пусть V - многообразие ассоциативных Z_2 -градуированных алгебр над

полем нулевой характеристики. Тогда следующие условия эквивалентны:

- (i) многообразие имеет конечную кодлину;
- (ii) многообразие имеет полиномиальный рост.

Доказательство.

Докажем, что из условия (i) следует условие (ii).

Пусть многообразие V имеет конечную кодлину. Поскольку кодлины многообразий $\text{supvar}G$, $\text{supvar}UT_2$, $\text{supvar}G^{gr}$, $\text{supvar}UT_2^{gr}$ и $\text{supvar}(K \oplus tK)$, $t^2 = 1$ имеют полиномиальный рост (см Замечание 1), то алгебры, порождающие данные многообразия не содержатся в многообразии V . Отсюда ([6]), мы получаем, что рост многообразия V полиномиален.

Теперь докажем, что из условия (ii) следует условие (i).

Пусть V многообразие ассоциативных Z_2 -градуированных алгебр полиномиального роста и его n -й характер равен

$$\chi_n^{gr}(V) = \sum_{r=0}^n \sum_{\substack{\lambda \vdash r \\ \mu \vdash n-r}} m_{\lambda, \mu} \chi_{\lambda, \mu} \quad (2)$$

Из полиномиальности роста ([4]) следует, что существует такая константа s , что все $S_r \times S_{n-r}$ -модули, $r = 0, \dots, n$, соответствующие парам $(r, n-r)$ -разбиений (λ, μ) , $|\lambda| - \lambda_1 > s$ или $|\mu| > s$, являются нулевыми. Иными словами, при любом n количество ненулевых неизоморфных подмодулей в разложении модуля $P_n^{gr}(V)$ на неприводимые не превосходит некоторой константы $P = (p(1) + \dots + p(s))^2$, не зависящей от n , где $p(i)$ число различных разбиений числа i .

Для доказательства конечности кодлины нам достаточно доказать, что существует такая константа C , не зависящая от n , что выполняется условие $m_{\lambda, \mu}(V) < C$ для всех n и всех пар разбиений (λ, μ) из суммы (2).

Возьмем $C = (2s)!(2s+1)(N+1)^{2s}$, N — степень тождества (1). Предположим, что существует пара $(r, n-r)$ -разбиений (λ, μ) , $\lambda = (\lambda_1, \dots, \lambda_{l_1})$, $\mu = (\mu_1, \dots, \mu_{l_2})$, такая что соответствующий ему модуль $W_{\lambda, \mu}$ входит в разложения модуля P_n^{gr} на неприводимые с кратностью $m_{\lambda, \mu} = m > C$. Пусть

$$W_{\lambda, \mu} = KH_n f_1 \oplus \dots \oplus KH_n f_m. \quad (3)$$

где $f_1, \dots, f_m$ -идемпотенты, построенные по разным парам таблиц Юнга (D_λ, D_μ) . Согласно ([6]) каждый подмодуль модуля $W_{\lambda, \mu}$ порождается элементом, который состоит из r четных переменных y_i , $y_i^\phi = y_i$, $i = 1, \dots, r$ и $n-r$, $n-r = |\mu| < s$ нечетных переменных z_i , $z_i^\phi = -z_i$, $i = 1, \dots, n-r$. В подмодулях $KH_n f_i$ существуют элементы $g_i \in KH_n f_i$, $i = 1, \dots, m$ с симметрическим набором от переменных $y_1, \dots, y_{l_1}$.

В этих элементах отождествим переменные указанного симметрического набора с переменной y . Полученные элементы обозначим через g'_i , $i = 1, \dots, m$ и докажем, что эти элементы будут линейно зависимыми. Для этого рассмотрим пространство

$$U = \langle y^{\alpha_0} u_{i_1} y^{\alpha_1} u_{i_2} \dots y^{\alpha_{k-1}} u_{i_k} y^{\alpha_k} \rangle,$$

где $k = n - \lambda_1 = |\lambda| + |\mu| - \lambda_1 \leq 2s$, $\alpha_0 + \dots + \alpha_k = \lambda_1$, а среди переменных u_{i_j} , $j = 1, \dots, k$ встречается $r - \lambda_1$ переменных y_i , $\lambda_1 + 1 \leq i \leq r$ и $n-r$ переменных z_i , $i = 1, \dots, n-r$. Оценим размерность данного пространства.

Нам дано, что рост нашего многообразия V полиномиальный, а так как рост многообразия $\text{supvar}UT_2$ почти полиномиальный [6], то выполняется условие $UT_2 \notin V$, и в V по Лемме 4.1 будет выполнено тождество (1). Применяя данное тождество и тождество ассоциативности к произвольному элементу пространства U , мы получим, что исходный элемент линейно выражается через элементы вида:

$$y^{\alpha'_0} u_{i_1} y^{\alpha'_1} u_{i_2} \dots y^{\alpha'_{k-1}} u_{i_k} y^{\alpha'_k}, \quad (4)$$

где среди чисел α'_i , $i = 0, \dots, k$ существует только одно $\alpha_j \geq N$, а остальные $\alpha_i < N$. Оценим число таких элементов.

Для удобства, слова вида $y^{\alpha'_i}$, $i = 0, \dots, k$ будем называть "зазорами", и если $\alpha_i \geq N$ то соответствующий "зазор" будем называть большим, в противном случае - маленьким. Число видов маленьких "зазоров" определяется числом переменных y в них, поэтому не превосходит числа $N + 1$. Всего маленьких зазоров в элементах вида (4) k штук. Отсюда, если число элементов вида (4) с фиксированным местоположением большого зазора не превосходит числа $(N + 1)^k$. Учитывая всевозможные перестановки переменных u_i , $j = 1, \dots, k$ и тот факт, что большой зазор может быть на одной из $k + 1$ позиций, то общее количество элементов вида (4) не превосходит числа $k!(k + 1)(N + 1)^k \leq (2s)!(2s + 1)(N + 1)^{2s} = C$. Мы получили, что $\dim U \leq C$.

Таким образом, среди элементов g'_i , $i = 1, \dots, m$, $m \geq C$ будут линейно зависимые. А поскольку мы работаем в случае поля нулевой характеристики, то среди элементов g_i , $i = 1, \dots, m$, $m \geq C$ также будут линейно зависимые. Что противоречит тому, что элементы g_i были выбраны из различных прямых слагаемых в сумме (3). Таким образом, число слагаемых в разложении (3) ограничено константой C .

То есть, в разложении модуля P_n^{gr} на неприводимые участвует не более P подмодулей с кратностями, ограниченными константой C , где числа P и C не зависят от n .

Теорема доказана.

Автор очень признателен С.П. Мищенко за всестороннюю помощь и внимание к работе.

Работа частично поддержана грантом "Университеты России - фундаментальные исследования" 1998 года.

Литература

- [1] Мальцев А.И. Алгебраические системы. - М: Наука, 1970.
- [2] Мищенко С.П. О многообразиях алгебр Ли промежуточного роста // Весті АН БССР.- 1987.- N 2.- С. 42 - 45.

- [3] *Drensky V. and Giambruno A.* Cocharacters, codimensions and Hilbert series of the polynomial identities for 2×2 matrices with involution, *Canadian J. Math.* 46 (1994), pp. 718-733.
- [4] *Giambruno A. and Mishchenko S.* Polynomial growth of the $*$ -codimensions and Young diagrams, (preprint).
- [5] *Giambruno A., Mishchenko S., M. Zaicev* Group actions and asymptotic behaviour of graded polynomial identities., to appear.
- [6] *Giambruno A., Mishchenko S., Zaicev M.* Superalgebras with almost polynomial growth, (preprint).
- [7] *Kemer A.* T-ideals with power growth of the codimensions are Specht, *Sibirskii Matematicheskii Zhurnal* 19 (1978), C. 54-69 (Russian); English translation: *Siberian Math. J.* 19 (1978), pp. 37-48.
- [8] *Mishchenko S., Regev A., Zaicev M.* A Characterization of P.I. Algebras wuth Bounded Multiplicities of the Cocharacters // *Journal of Algebra*, 1999, v.219, No 1, pp. 356-368

АСИМПТОТИКА РЕШЕНИЙ ЗАДАЧИ ОБ УСТОЙЧИВОСТИ ПЛАСТИНЫ В СВЕРХЗВУКОВОМ ПОТОКЕ ГАЗА.

П.А. Вельмисов, С.В. Киреев, А.О. Кузнецов.

Получены условия статической неустойчивости (бифуркации) и динамической устойчивости пластины-полосы в сверхзвуковом потоке газа.

1. Рассмотрим двухточечную задачу об изгибных формах пластины-полосы в сверхзвуковом потоке газа [1,2], описываемые нелинейным уравнением :

$$L(w) \equiv K(w) + \sum_{n=1}^{\infty} a_{2n+1} w^{2n+1} - \theta w'' \int_0^{\ell} w'^2 dx = 0 \quad (1)$$

$$K(w) \equiv Dw'''' + \alpha^* w', \quad \alpha^* = \frac{\alpha_0 \rho_0 V^2}{\sqrt{M^2 - 1}}, \quad M = \frac{V}{a}$$

Здесь D - изгибная жесткость пластины; V , ρ_0 , a - скорость газа, плотность и скорость звука, соответствующие однородному потоку; M - число Маха; a_j - коэффициенты, характеризующие жесткость основания; интегральный член учитывает нелинейное воздействие продольного усилия; $\alpha w'$ - член, учитывающий аэродинамическое воздействие; $\alpha_0 = 1$ ($\alpha_0 = 2$) соответствует одностороннему (двустороннему) обтеканию пластины; $w(x)$ - прогиб пластины. Все коэффициенты, входящие в уравнение, постоянные.

Для поиска малых решений задачи (1), ответвляющихся от нулевого, применим методы теории ветвления [3].

Подробно рассмотрим уравнение (1), оставляя только первый член ряда :

$$Dw'''' + \alpha w' + a_3 w^3 - \theta w'' \int_0^{\ell} w'^2 dx = 0$$

(2)

Перейдем в уравнении (2) к безразмерным переменным. После замены

$$x = \ell \bar{x}, \quad w = \ell \bar{w}$$

(3)

где ℓ - некоторый характерный размер, а величины с чертой - безразмерные переменные, приходим к уравнению

$$\bar{w}'''' + \frac{\alpha \ell^3}{D} \bar{w}' + \frac{a_3 \ell^6}{D} \bar{w}^3 - \frac{\theta \ell^3}{D} \bar{w}'' \int_0^{\ell} \bar{w}'^2 d\bar{x} = 0 \quad (4)$$

Будем изучать решения этого уравнения при граничных условиях

- 1) $\bar{w}''(0) = 0, \bar{w}'''(0) = 0; \quad \bar{w}(1) = 0, \bar{w}'(1) = 0,$
- 2) $\bar{w}'(0) = 0, \bar{w}'''(0) = 0; \quad \bar{w}(1) = 0, \bar{w}''(1) = 0,$
- 3) $\bar{w}'(0) = 0, \bar{w}'''(0) = 0; \quad \bar{w}(1) = 0, \bar{w}'(1) = 0,$

(5)

соответствующих :

- 1) свободному и жестко заземленному концам,
- 2) заземлению в свободный элемент и шарнирному закреплению,
- 3) заземлению в свободный элемент и жесткому заземлению.

В дальнейшем изложении для упрощения записи мы будем опускать черточки, но понимая все величины как безразмерные.

Линеаризованная система

$$K(w) = w'''' + \lambda \cdot w', \quad \lambda = \frac{\alpha \cdot \ell^3}{D}$$

(7)

- 1) $w''(0) = 0, w'''(0) = 0; \quad w(1) = 0, w'(1) = 0,$
- 2) $w'(0) = 0, w'''(0) = 0; \quad w(1) = 0, w''(1) = 0,$
- 3) $w'(0) = 0, w'''(0) = 0; \quad w(1) = 0, w'(1) = 0,$

определяет фредгольмов несамоспряженный оператор B , действующий из пространства $C^{4+\alpha}[0,1]$ в пространство $C^\alpha[0,1]$.

Общее решение уравнения (7) имеет вид :

$$w(x) = c_1 + c_2 e^{-sx} + e^{\frac{sx}{2}} \left(c_3 \cos\left(\frac{\sqrt{3}}{2} sx\right) + c_4 \sin\left(\frac{\sqrt{3}}{2} sx\right) \right) \quad (8)$$

где c_1, c_2, c_3, c_4 - константы, а величина s определяется из соотношения :

$$s^3 = \frac{\alpha \cdot \ell^3}{D}$$

(9)

Собственные числа $\lambda = s^3$ оператора B определяются из соответствующих каждому граничному условию (5) дисперсионных соотношений :

$$\begin{aligned} 1), 2) \quad & \frac{1}{2} e^{-\frac{3}{2}s} + \cos\left(\frac{\sqrt{3}}{2} s\right) = 0, \\ 3) \quad & \cos\left(\frac{\sqrt{3}}{2} s - \frac{\pi}{3}\right) - \frac{1}{2} e^{-\frac{3}{2}s} = 0 \end{aligned} \quad (10)$$

Собственным числам из (10) отвечают собственные функции

$$\begin{aligned} 1) \quad \varphi &= 2\sqrt{3} \sin\left(\frac{\sqrt{3}}{2} s + \frac{\pi}{3}\right) e^{\frac{s}{2}} + e^{-sx} - 2e^{\frac{s}{2}x} \sin\left(\frac{\sqrt{3}}{2} sx + \frac{\pi}{6}\right), \\ 2), 3) \quad \varphi &= -2\sqrt{3} \sin\left(\frac{\sqrt{3}}{2} s - \frac{\pi}{3}\right) e^{\frac{s}{2}} + e^{-sx} + 2e^{\frac{s}{2}x} \sin\left(\frac{\sqrt{3}}{2} sx - \frac{\pi}{6}\right) \end{aligned} \quad (11)$$

Следуя [4], получим сопряженные задачи из условия

$$\int_0^1 u(w'''' + \lambda \cdot w') dx = 0$$

для каждого граничного условия (5) соответственно :

$$1) K^*(u) = u'''' - \lambda \cdot u' = 0,$$

$$u''(0) = 0, \quad u'''(0) - \lambda \cdot u(0) = 0, \quad u(1) = 0, \quad u'(1) = 0,$$

$$2) K^*(u) = u'''' - \lambda \cdot u' = 0,$$

$$u'(0) = 0, \quad u'''(0) - \lambda \cdot u(0) = 0, \quad u(1) = 0, \quad u''(1) = 0,$$

$$3) K^*(u) = u'''' - \lambda \cdot u' = 0,$$

$$u'(0) = 0, \quad u'''(0) - \lambda \cdot u(0) = 0, \quad u(1) = 0, \quad u'(1) = 0$$

Они имеют те же собственные числа (10) при соответствующих собственных функциях:

$$\begin{aligned} 1) \psi(x) &= e^x + \left(\frac{1}{2} + \sqrt{3} e^{\frac{3}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) \right) e^{\frac{s}{2}x} \cos\left(\frac{\sqrt{3}}{2}sx\right) + \left(\frac{\sqrt{3}}{2} - e^{\frac{3}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) \right) e^{\frac{s}{2}x} \sin\left(\frac{\sqrt{3}}{2}sx\right), \\ 2) \psi(x) &= -e^x + \left(-\frac{1}{2} + \sqrt{3} e^{\frac{3}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) \right) e^{\frac{s}{2}x} \cos\left(\frac{\sqrt{3}}{2}sx\right) + \left(\frac{\sqrt{3}}{2} + e^{\frac{3}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) \right) e^{\frac{s}{2}x} \sin\left(\frac{\sqrt{3}}{2}sx\right), \\ 3) \psi(x) &= -\frac{1}{2}e^x + \left(\frac{1}{2} - \sqrt{3} e^{\frac{3}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) \right) e^{\frac{s}{2}x} \cos\left(\frac{\sqrt{3}}{2}sx\right) + \left(\frac{\sqrt{3}}{2} - e^{\frac{3}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) \right) e^{\frac{s}{2}x} \sin\left(\frac{\sqrt{3}}{2}sx\right). \end{aligned} \quad (12)$$

Полагая $\varepsilon = \lambda - s_0^3$ (s_0^3 - точка бифуркации) и применяя лемму Шмидта [3], запишем отвечающее (7) уравнение (4) в виде системы :

$$\tilde{K}(w) = w'''' + s^3 w' + \langle w, \gamma \rangle z = \xi z - \varepsilon w' + \frac{\theta \ell^3}{D} w'' \int_0^1 w'^2 dx - \frac{a_3 \ell^6}{D} w^3 \quad (13)$$

$$\xi = \langle w, \gamma \rangle,$$

где $\{\varphi, \gamma\}$ и $\{z, \psi\}$ - некоторые биортогональные системы относительно внешнего произведения

$$\langle u, v \rangle = \int_0^1 u(x)v(x)dx$$

Разыскивая решение первого уравнения в виде ряда $w = \sum_{k+j \geq 1} w_{kj} \xi^k \varepsilon^j$

($w_{0j} = 0$, т.к. (1) - задача о точке бифуркации), получаем рекуррентную систему для определения w_{kj} :

$$w_{10} = \varphi, \quad \tilde{K}w_{11} = -\varphi', \quad \tilde{K}w_{30} = -\frac{a_3 \ell^6}{D} \varphi^3 + \frac{\theta \ell^3}{D} \varphi'' \int_0^1 \varphi'^2 dx, \quad \dots$$

Тогда второе уравнение системы представляет собой уравнение разветвления :

$$L_{11}\xi\varepsilon + L_{30}\xi^3 + \dots = 0$$

(14)

и дает асимптотику разветвляющихся решений

$$w(x) = \pm \sqrt{-\frac{L_{11}\varepsilon}{L_{30}}\varphi + o\left(|\varepsilon|^{\frac{1}{2}}\right)} \quad (15)$$

$$\text{sign } \varepsilon = -\text{sign } L_{11} \cdot L_{30}$$

с коэффициентами,

$$L_{11} = -\langle \varphi', \psi \rangle = -\int_0^1 \varphi' \psi dx$$

$$L_{30} = \langle -\frac{a_3 \ell^6}{D} \varphi^3 + \frac{\theta^3}{D} \varphi'' \int_0^1 \varphi'^2 dx, \psi \rangle = -\frac{a_3 \ell^6}{D} \langle \varphi^3, \psi \rangle + \frac{\theta^3}{D} \langle \varphi'' \int_0^1 \varphi'^2 dx, \psi \rangle$$

$$= -\frac{a_3 \ell^6}{D} \int_0^1 \varphi^3 \psi + \frac{\theta^3}{D} \int_0^1 \varphi'' \psi \int_0^1 \varphi'^2 dx = -\frac{a_3 \ell^6}{D} L_{30}^{(1)} + \frac{\theta^3}{D} L_{30}^{(2)}$$

Коэффициенты L_{11} , $L_{30}^{(1)}$ и $L_{30}^{(2)}$ соответственно для каждого типа закрепления (5) имеют вид:

$$1) \quad L_{11} = \frac{3}{2}s + \sqrt{3}s \cdot e^{\frac{3}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right),$$

$$L_{30}^{(1)} = \frac{1523}{728 \cdot s} - \frac{9595}{364 \cdot s} e^s + \frac{279}{28 \cdot s} e^{-s} - \frac{126}{7 \cdot s} e^{2s} - \frac{2485}{364 \cdot s} e^{-2s} + \frac{847}{728 \cdot s} e^{-5s} + \frac{15\sqrt{3}}{14 \cdot s} e^{\frac{s}{2}} \sin\left(\frac{\sqrt{3}}{2}s\right) -$$

$$- \frac{355\sqrt{3}}{364 \cdot s} e^{\frac{s}{2}} \sin\left(\frac{\sqrt{3}}{2}s\right) - \frac{22\sqrt{3}}{13 \cdot s} e^{\frac{3}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) + \frac{6929\sqrt{3}}{2366 \cdot s} e^{\frac{5}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) + \frac{9\sqrt{3}}{2} e^{\frac{5}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right) +$$

$$+ \frac{47\sqrt{3}}{104 \cdot s} e^{\frac{7}{2}s} \sin\left(\frac{\sqrt{3}}{2}s\right)$$

$$L_{30}^{(2)} = \frac{9}{4}s^3e^{-2s} + 3\sqrt{3}s^2e^{-\frac{s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right)\left(2+\frac{1}{2}s\right) + \frac{9}{2}s^2e^s\left(3-\frac{1}{2}s\right) + 3\sqrt{3}s^2e^{\frac{5s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right)\left(1-\frac{1}{2}s\right)$$

$$2) \quad L_{11} = -\frac{3}{2}s - \sqrt{3}s \cdot e^{\frac{3s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right),$$

$$L_{30}^{(2)} = \frac{9}{2}s^2e^s(1-s) - \frac{27}{4}s^2(1-s) - 3\sqrt{3}s^2e^{-\frac{s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right)(1+s) + 3\sqrt{3}s^2e^{\frac{5s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right)(1+s) -$$

$$- \frac{9\sqrt{3}}{2}s^2e^{\frac{3s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right)(1+s) + \frac{1}{2}s^2e^{-2s}\left(9s - \frac{33}{4}\right) - \frac{3}{8}s^2e^{-5s} - \frac{9}{16}se^{-3s}$$

$$L_{30}^{(1)} = -\frac{732069}{37856 \cdot s} + \frac{918}{91 \cdot s}e^s - \frac{425}{56 \cdot s}e^{-s} + \frac{483}{56 \cdot s}e^{2s} - \frac{2325}{2912 \cdot s}e^{-2s} + \frac{119}{26 \cdot s}e^{-3s} + \frac{63\sqrt{3}}{56 \cdot s}e^{\frac{s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right) +$$

$$+ \frac{33179\sqrt{3}}{1456 \cdot s}e^{-\frac{s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right) - \frac{1395\sqrt{3}}{91 \cdot s}e^{\frac{3s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right) + \frac{534\sqrt{3}}{91 \cdot s}e^{\frac{5s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right) - \frac{9\sqrt{3}}{s}e^{-\frac{3s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right) +$$

$$+ \frac{100\sqrt{3}}{91 \cdot s}e^{-\frac{7s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right) - \frac{9}{s}e^{3s} + \frac{351}{728 \cdot s}e^{-5s}$$

$$3) \quad L_{11} = -\frac{3}{2}s + \sqrt{3}s e^{\frac{3s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right) - \frac{\sqrt{3}}{6}e^{\frac{3s}{2}}\sin\left(\frac{\sqrt{3}}{2}s\right),$$

$$L_{30}^{(1)} = \frac{147455}{2548 \cdot s} + \frac{29650}{728 \cdot s}e^s + \frac{1073}{112 \cdot s}e^{-2s} - \frac{27}{2 \cdot s}e^{-3s} + \frac{56}{13 \cdot s}e^{-5s} + \left(\left(\frac{39763\sqrt{3}}{1456 \cdot s} - \frac{27}{7 \cdot s}\right)e^{-\frac{s}{2}} +$$

$$+ \frac{3639\sqrt{3}}{182 \cdot s}e^{\frac{5s}{2}} - \frac{10071\sqrt{3}}{91 \cdot s}e^{\frac{3s}{2}} - \frac{1447\sqrt{3}}{91 \cdot s}e^{-\frac{7s}{2}} + \frac{27\sqrt{3}}{s}e^{-\frac{3s}{2}} - \frac{5\sqrt{3}}{14 \cdot s}e^{\frac{s}{2}}\right)\sin\left(\frac{\sqrt{3}}{2}s\right) +$$

$$+ \left(\left(\frac{54\sqrt{3}}{7 \cdot s} - \frac{1795}{728 \cdot s}\right)e^s + \frac{4204}{91 \cdot s}e^{-2s} + \frac{144}{s}e^{3s}\right)\sin\left(\frac{\sqrt{3}}{2}s\right)^2 + \left(\left(\frac{583\sqrt{3}}{182 \cdot s} - \frac{81}{7 \cdot s}\right)e^{\frac{5s}{2}} +$$

$$+ \left(\frac{1530\sqrt{3}}{91 \cdot s} + \frac{108}{7 \cdot s}\right)e^{-\frac{s}{2}} - \frac{3\sqrt{3}}{s}e^{-\frac{3s}{2}}\right)\sin\left(\frac{\sqrt{3}}{2}s\right)^3 + \left(\frac{9}{s} - \left(-\frac{216\sqrt{3}}{7 \cdot s} + \frac{2773}{91 \cdot s}\right)e^s\right)\sin\left(\frac{\sqrt{3}}{2}s\right)^4 +$$

$$+ \sin\left(\frac{\sqrt{3}}{2}s\right)^5 \left(-\frac{1340\sqrt{3}}{91 \cdot s} + \frac{324}{7 \cdot s}\right)e^{\frac{5s}{2}}$$

$$\begin{aligned}
 L_{30}^{(2)} = & -\frac{27}{2}s^2e^s - \frac{81}{4}s^2 + \frac{33}{4}s^2e^{-2s} + \frac{9}{4}s^2e^{-3s} - \frac{3}{4}s^2e^{-3s} - 24s^2\cos\left(\frac{\sqrt{3}}{2}s\right)^4 + 18s^2\cos\left(\frac{\sqrt{3}}{2}s\right)^2 + \\
 & + 39s^2e^s\cos\left(\frac{\sqrt{3}}{2}s\right)^2 - 9s^2e^{-2s}\cos\left(\frac{\sqrt{3}}{2}s\right)^2 - 18\sqrt{3}s^2e^{2s}\sin\left(\frac{\sqrt{3}}{2}s\right)\cos\left(\frac{\sqrt{3}}{2}s\right)^2 + \\
 & + 24\sqrt{3}s^2e^{2s}\sin\left(\frac{\sqrt{3}}{2}s\right)\cos\left(\frac{\sqrt{3}}{2}s\right)^4 + 6\sqrt{3}s^2e^{-2s}\sin\left(\frac{\sqrt{3}}{2}s\right)\cos\left(\frac{\sqrt{3}}{2}s\right)^2 - \frac{3\sqrt{3}}{2}s^2e^{-2s}\sin\left(\frac{\sqrt{3}}{2}s\right)(3+s) - \\
 & - 6\sqrt{3}s^2e^{2s}\sin\left(\frac{\sqrt{3}}{2}s\right)\cos\left(\frac{\sqrt{3}}{2}s\right)^2(5+s) + 3\sqrt{3}s^2e^{2s}\sin\left(\frac{\sqrt{3}}{2}s\right)(3+s) + \frac{9\sqrt{3}}{2}s^2e^{2s}\sin\left(\frac{\sqrt{3}}{2}s\right)(3+s)
 \end{aligned}$$

Для полученных асимптотических решений (15) построены бифуркационные диаграммы, показывающие зависимость максимального прогиба стержня от скорости набегающего потока ($V \gg a$). Если возмущение $\varepsilon = \lambda - s_0^3$ образуется за счет изменения скорости потока, то ветвление в точке s_0^3 надкритическое, то есть $-L_{11}/L_{30} > 0$, откуда $\varepsilon(O(V)) > 0$ образуется увеличением скорости потока. В то же время $\varepsilon(1/D) > 0$, при неизменной скорости потока, может быть получено вследствие уменьшения изгибной жесткости, для которой ветвление будет подкритическим. На рис. 1, рис. 2, представлены бифуркационные диаграммы прогиба при фиксированных коэффициентах изгибной жесткости $D_1 < D_2 < D_3$ в зависимости от изменения скорости набегающего потока сверх критических значений $\lambda_1 < \lambda_2 = s_0^3 < \lambda_3$. Рисунок 1 соответствует первому, а рисунок 2 - третьему типу закреплений (5).

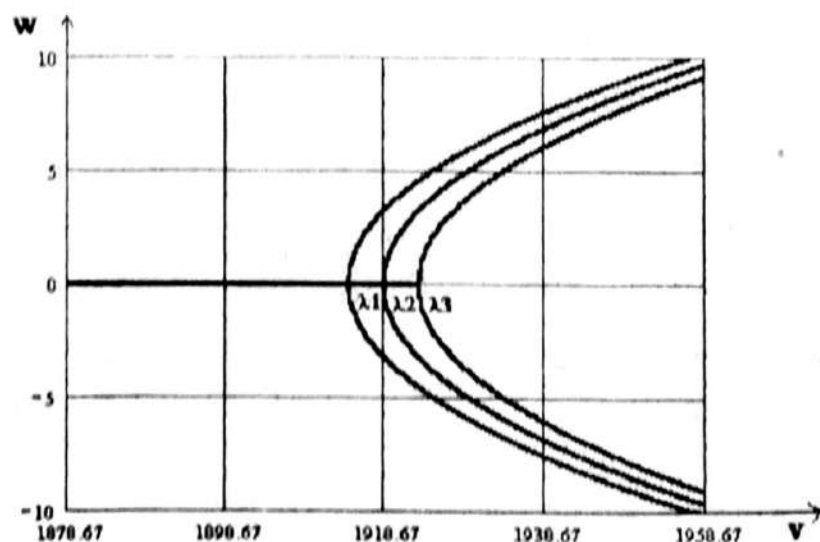


рис. 1

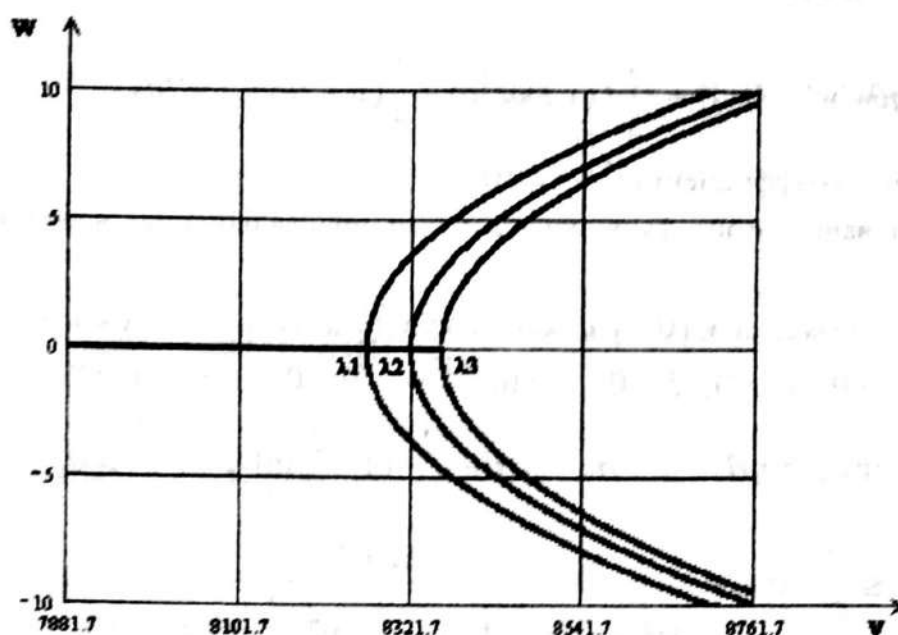


рис.2

Здесь : $a_3 = 1$, $\theta = 1$, $\ell = 10$, $\rho_0 = 1$, $a = 330$;

на рис.1 : $D_1 = 10^{7.999}$, $D_2 = 10^8$, $D_3 = 10^{8.001}$,

на рис.2 : $D_1 = 10^{7.997}$, $D_2 = 10^8$, $D_3 = 10^{8.002}$.

2. Для описания колебаний пластины в сверхзвуковом потоке газа при $V \gg a$ будем использовать уравнение, являющееся обобщением уравнения (1)

$$m\ddot{w} + (Dw'')'' + (\xi\dot{w}'')'' - (\eta\dot{w}')' + (Nw')' + \gamma\dot{w} + \beta w + \\ + \alpha_*(\dot{w} + Vw') + f(x, w) = w'' \left(\mu \int_0^\ell w'^2 dx + \nu \frac{\partial}{\partial t} \int_0^\ell w'^2 dx \right) \quad (16)$$

где $m(x)$ - удельная масса; $D(x)$ - изгибная жесткость; $\xi(x), \gamma(x)$ - коэффициенты внутреннего и внешнего демпфирования; $\eta(x)$ - коэффициент, учитывающий инерцию вращения; $\beta(x)$ - коэффициент жесткости основания; $N(x)$ - коэффициент, учитывающий вес пластины, а также сжимающее (растягивающее) усилие; $\alpha_* = \alpha_0 \rho_0 a$; μ, ν - коэффициенты, учитывающие степень ограничения на перемещение концов и возникающее при этом демпфирование ($\mu, \nu = const$); $f(x, w)$ - нелинейная составляющая реакции основания; точка и штрих сверху соответствуют частным производным по времени t и координате x .

Введем в рассмотрение функционал

$$J(t) = \frac{1}{2} \int_0^\ell \{ m\dot{w}^2 + 2M\theta w\dot{w} + (\beta + \gamma\theta + \alpha_*\theta)w^2 + (D + \xi\theta)w''^2 + \eta\dot{w}'^2 +$$

$$+ 2\eta\theta w'\dot{w}' - Nw'^2 + 2\int_0^w f(x,z)dz \Big\} dx + \frac{1}{4}(\mu + 2\theta\nu) \left(\int_0^\ell w'^2 dx \right)^2 \quad (17)$$

где θ - некоторый коэффициент ($\theta = \text{const}$).

Проводя исследование этого функционала и его производной по времени, можно доказать

Теорема Пусть для всех $x \in [0, \ell]$ и любого $w(x, t)$ выполнены неравенства

$$M > 0, D > 0, \xi \geq 0, \eta \geq 0, \beta \geq 0, \theta > 0, \mu \geq 0, \nu \geq 0$$

$$\frac{2}{\ell^2} \inf_x \xi \geq \eta\theta, \quad \gamma + \alpha_* - M\theta + \frac{2}{\ell^2} \inf_x \left(\frac{2}{\ell^2} \inf_x \xi - \eta\theta \right) > 0,$$

$$N \leq \frac{2}{\ell^2} \inf_x D - \frac{\alpha_*^2 V^2}{4\theta \left[\gamma + \alpha_* - M\theta + \frac{2}{\ell^2} \inf_x \left(\frac{2}{\ell^2} \inf_x \xi - \eta\theta \right) \right]}$$

$$wf(x, w) \geq 0, \quad \int_0^w f(x, z) dz \geq 0$$

Тогда решение $\{w(x, t), \dot{w}(x, t)\}$ уравнения (16) устойчиво (по Ляпунову) в среднем на $[0, \ell]$ по отношению к возмущениям начальных значений $w''(x, 0), \dot{w}'(x, 0)$, если выполнены следующие граничные условия: "ж-ж", "ж-ш", "ш-ж", "ш-ш", "ж-с.э." (ж - жесткое защемление, ш - шарнирное закрепление, с.э. - защемление в свободный элемент).

Область устойчивости на плоскости (V, N) ограничена ветвью параболы, направленной вниз.

3. Деформации пластины-полосы, предварительно сжатой продольным усилием $N(N = \text{const})$, описывается уравнением

$$Dw'''' + Nw'' + \alpha w' - \theta w'' \int_0^\ell w'^2(x) dx = 0 \quad (18)$$

Сформулируем теорему, позволяющую находить точное решение краевых задач для этого уравнения

Теорема Если $w_*(x)$ - решение некоторой линейной краевой задачи для уравнения $Dw_*'''' + \alpha w_*' = 0$, то

$$w(x) = \pm \sqrt{\frac{N}{\theta}} \frac{w_*}{\sqrt{\int_0^\ell w_*'^2(x) dx}}$$

- точное решение нелинейной краевой задачи для уравнения (18) с соответствующими граничными условиями.

ЛИТЕРАТУРА

1. Вольмир А.С. *Нелинейная динамика пластин и оболочек*. - М.: Наука, 1972. - 432с.
2. Пановко Я.Г., Губанова И.И. *Устойчивость и колебания упругих систем*. - М.: Наука, 1979. - 384с.
3. Вельмисов П.А., Логинов Б.В. *Метод групповых преобразований и ветвление решений в двухточечных граничных задачах аэроупругости*. // Материалы международной конференции "Дифференциальные уравнения и их приложения", 1995, издательство Мордовского университета, с.120.
4. Найфэ А.Х. *Методы возмущений*. - М.: Мир, 1976. - 456с.
5. Вайнберг М.М., Треногин В.А. *Теория ветвления решений нелинейных уравнений*. - М.: Наука, 1969. - 524с.

U^* — сопряженный оператор. Векторное пространство $\bar{H}$ с введенным псевдоскалярным произведением называется псевдогильбертовым пространством. Элементы $\bar{H}$ называются псевдовекторами, линейные непрерывные операторы, действующие в H_0 — псевдоскалярами. Условимся впредь обозначать через $x, y, \dots$ — векторы из H_0 , через $f, g, \dots$ — векторы из H , а через $U, V, P, \dots$ — псевдовекторы из $\bar{H}$.

Напомним свойства псевдоскалярного произведения; для любых псевдовекторов U, V, P :

1. $\{U, V + P\} = \{U, V\} + \{U, P\}$;
2. $\{U, VC\} = \{U, V\}C$; $\{UC, V\} = C^*\{U, V\}$, где C — псевдоскаляр;
3. $\{U, V\}^* = \{V, U\}$.

Всякий линейный оператор T , действующий в H с областью определения D_T , естественным образом порождает в псевдогильбертовом пространстве $\bar{H}$ оператор $\bar{T}$:

1. $D_{\bar{T}} = \{U \in \bar{H} : Ux \in D_T \forall x \in H_0\}$, где $D_{\bar{T}}$ — область определения оператора $\bar{T}$;
2. $(\bar{T}U)x = T(Ux) \forall U \in D_{\bar{T}}, x \in H_0$.

Очевидно, что при любом $x \in H_0, x \neq 0, D_{\bar{T}} = \{f \in H : f = Ux, U \in D_{\bar{T}}\}$.

Всюду в дальнейшем под $\bar{T}$ будем понимать оператор в $\bar{H}$, порожденный оператором T . Элементы $U, V \in \bar{H}$ называются псевдоортогональными, если $\{U, V\} = 0$, где 0 — нулевой оператор в H_0 .

2. Формула обобщенных резольвент изометрических операторов

Пусть V — произвольный изометрический оператор в гильбертовом пространстве H с произвольными дефектными числами.

Будем устанавливать формулу обобщенных резольвент оператора V в терминах псевдогильбертовых пространств.

Через D_V и Δ_V будем обозначать соответственно область определения и область значений оператора V .

При любом комплексном $\zeta (|\zeta| \neq 1)$ положим:

$$M_\zeta = (I - \zeta V)D_V, \quad N_\zeta = H \ominus M_\zeta,$$

где N_ζ — дефектное подпространство оператора V отвечающее комплексному числу ζ .

В частности, при $\zeta = 0$ имеем:

$$M_0 = D_V, \quad N_0 = H \ominus D_V; \quad M_\infty = \Delta_V, \quad N_\infty = H \ominus \Delta_V.$$

Для определенности будем считать

$$\dim N_0 \geq \dim N_\infty.$$

ОБ ОДНОЙ ФОРМУЛЕ ОБОБЩЕННЫХ РЕЗОЛЬВЕНТ ИЗОМЕТРИЧЕСКИХ ОПЕРАТОРОВ

Г.М. Ильмушкин

Впервые обобщенные резольвенты замкнутого симметрического плотно заданного оператора с дефектными числами, равными единице, были описаны в 1943 году независимо друг от друга М.Г.Крейном [1] и М.А.Наймарком [2]. Впоследствии М.Г.Крейн [3] обобщил свою формулу на случай любых конечных равных дефектных чисел. На случай бесконечных равных дефектных чисел формула М.Г.Крейна была перенесена Ш.Н.Саакяном [6]. В ряде своих работ [4,5,7] А.В.Штраус получил иную формулу, описывающую обобщенные резольвенты симметрических операторов с произвольными дефектными числами. При этом формула А.В.Штрауса охватывает и случай операторов с неплотной областью определения. На случай изометрического оператора формула обобщенных резольвент А.В.Штрауса обобщена М.Е.Чумакиным [8]. В дальнейшем в исследованиях [10—13] были установлены формулы обобщенных резольвент неплотно заданных симметрических и изометрических операторов с произвольными дефектными числами через фиксированную резольвенту.

В предлагаемой работе устанавливается формула обобщенных резольвент изометрического оператора с произвольными дефектными числами в терминах псевдогильбертовых пространств. Такой подход обусловлен тем, что иногда при построении обобщенных резольвент изометрического оператора естественным образом возникают псевдогильбертовы пространства; в частности, при исследовании операторных степенных проблем моментов, разностных уравнений с операторными коэффициентами [12,13]. Попутно в работе для изометрического оператора, не имеющего ненулевых неподвижных элементов, дается методом псевдогильбертовых пространств описание обобщенных резольвент, соответствующих унитарным расширениям без ненулевых неподвижных элементов.

§1. Предварительные сведения

При выводе формулы обобщенных резольвент изометрического оператора мы будем использовать методы псевдогильбертовых пространств [9]. Напомним некоторые определения и факты, связанные с этим понятием.

Пусть H_0 и H — гильбертовы пространства. Обозначим через $\overline{H} = L(H_0; H)$ пространство линейных непрерывных операторов, отображающих H_0 в H . Для любых $U, V \in \overline{H}$ введем псевдоскалярное произведение, полагая $\{U, V\} = U^*V$, где

Пусть H_0 — гильбертово пространство такое, что

$$\dim H_0 \geq \dim N_0.$$

$\overline{H} = L(H_0, H)$ как и раньше будет означать псевдогильбертово пространство.

Введем обозначения

$$\overline{M}_\zeta = (\overline{I} - \zeta \overline{V})D_V, \quad N_\zeta = \{U \in \overline{H} : \{U, V\} = 0, \forall V \in \overline{M}_\zeta\}.$$

Справедлива

Лемма 1. Имеет место следующее

$$\overline{N}_\zeta = \{U \in \overline{H} : Ux \in N_\zeta, \forall x \in N_0\}.$$

Запишем формулу обобщенных резольвент $R_\zeta(|\zeta| \leq 1)$ оператора V [8]:

$$R_\zeta = (I - \zeta U_{\Phi(\zeta)})^{-1}(|\zeta| \leq 1),$$

где $\Phi(\zeta)$ — произвольная регулярная в единичном круге $S = \{\zeta : |\zeta| \leq 1\}$ оператор-функция, значения которой суть линейные натягивающие операторы, отображающие N_0 в N_∞ , а $U_{\Phi(\zeta)}$ — расширение оператора V на пространство H , определяемое оператором $\Phi(\zeta)$ по формуле

$$\Phi(\zeta)(f + g) = Vf + \Phi(\zeta)g (f \in D_V, g \in N_0).$$

Напомним, что обобщенные резольвенты для $\zeta, |\zeta| > 1$ определяются из равенства

$$R_\zeta^* + R_{\frac{1}{\zeta}} = I.$$

Очевидно, что для любого $g \in N_\infty$:

$$U_{\Phi(\zeta)}^* g = \Phi^*(\zeta)g.$$

Обозначим через $K(S; \eta_1, \eta_2)$ класс регулярных в круге S оператор-функций $\Phi(\zeta)$, значениями которых являются линейные нерастягивающие операторы, отображающие одно гильбертово пространство η_1 в другое η_2 .

Пусть $R_\zeta^0(|\zeta| \neq 1)$ — произвольная фиксированная резольвента оператора V .

Отметим следующие свойства оператора R_ζ^0 :

1. $R_\zeta^0 M_\zeta = D_V$; 2. $R_\zeta^0 N_\infty = N_{\frac{1}{\zeta}}(|\zeta| \leq 1)$.

Обозначим через $U_\zeta^0 = R_\zeta^0|_{N_\infty}$ — сужение оператора R_ζ^0 на N_∞ .

Пусть I_0 — частично изометрический оператор, заданный на H_0 и отображающий $H_0 \ominus \text{Ker } I_\infty = H_I \subset H_0$ на N_0 , аналогично I_∞ — частично изометрический оператор, отображающий $H_0 \ominus \text{Ker } I_\infty = N_\infty \subset H_0$ на Δ_V .

Итак, $I_0 \in \overline{N}_0, I_\infty \in \overline{N}_\infty$.

Пусть нам известно семейство операторов $\{I_\zeta\}(\zeta \in S)$, аналитически зависящее от параметра ζ в круге S , и при каждом $\zeta \in S$: I_ζ — частично изометрический оператор, определенный на H_0 и отображающий на $N_{\frac{1}{\zeta}}$.

Для разности резольвент R_ζ и R_ζ^0 имеем

$$R_\zeta f - R_\zeta^0 f = \begin{cases} \in 0, & \text{если } f \in M_\zeta \\ \in N_{\frac{1}{\zeta}}, & \text{если } f \in N_\zeta \end{cases}$$

Отсюда для любых будем иметь

$$R_\zeta(Ux) = R_\zeta^0(Ux) + I_\zeta C(U, \zeta)x, \quad (1)$$

где $C(U, \zeta)$ — псевдоскалярная функция, которая надлежит определению.

Справедлива

Лемма 2. Для любых $U \in \overline{H}, P \in \overline{N}_\infty$:

$$\{P, \overline{U}_{\Phi(\zeta)} \overline{R}_\zeta U\} = \{\overline{\Phi}^*(\zeta) P \overline{R}_\zeta U\}. \quad (2)$$

Для определения $C(U, \zeta)$ в (1) воспользуемся леммой 2.

Пусть R_ζ — обобщенная резольвента, определяемая операторной функцией $\Phi(\zeta)$ класса $K(S; N_0, N_\infty)$. Подберем в (1) функцию $C(U, \zeta)$ так, чтобы выполнялось соотношение (2). Для этого достаточно взять $P = I_\infty$.

Итак, имеем

$$\{I_\infty, \overline{U}_{\Phi(\zeta)} [\overline{R}_\zeta^0 + I_\zeta C(U, \zeta)]\} = \{\overline{\Phi}^*(\zeta) I_\infty, \overline{R}_\zeta U + I_\infty C(U, \zeta)\} \quad (3)$$

Заметим, что произвольная функция $\Phi(\zeta)$ класса $K(S; N_0, N_\infty)$ определяется так

$$\Phi(\zeta) I_0 x = I_\infty w(\zeta)(\zeta)x, \text{ где } w(\zeta) \in K(S; H, H), x \in H_0 \quad (4)$$

Введем обозначения:

$$\Gamma(\infty, \zeta) = \{I_\infty, V_\zeta\}|_H, \quad \Gamma(0, \zeta) = \{I_0, V_\zeta\}|_H,$$

Сформулируем без доказательства следующее утверждение.

Теорема 1. Для любого $\zeta \in S$ существует ограниченный обратный оператор $\Gamma(\infty, \zeta)^{-1}$, отображающий H_∞ на H_ζ .

Далее введем в рассмотрение оператор-функцию $k(\zeta)$ параметра $\zeta \in S$ следующим образом

$$K(\zeta) = \zeta \Gamma(0, \zeta) \Gamma(\infty, \zeta)^{-1} P_{H_-},$$

где P_{H_-} — ортороектор в H_0 на H_∞ . Для $\frac{1}{\zeta}$ ($\zeta \in S$) эту функцию продолжим следующим образом

$$K\left(\frac{1}{\zeta}\right) = K^*(\zeta) \quad \forall \zeta \in S$$

Назовем эту функцию характеристической функцией оператора V .

Можно установить, что это определение связано с характеристической функцией $\chi(\zeta)$ в смысле определения, предложенного в [14]:

$$\chi(\zeta) = \zeta P(\zeta)|_{N_\infty}, \quad \text{если } \zeta \in S,$$

где $P(\zeta)$ — оператор проектирования на N_0 параллельно подпространству M_ζ и $\chi(\zeta) = \frac{1}{\zeta} P(\zeta)|_{N_0}$, если $|\zeta| > 1$, где $P(\zeta)$ — оператор проектирования на N_∞ параллельно M_ζ .

Оператор-функции $\chi(\zeta)$ и $K(\zeta)$ связаны следующим соотношением

$$\chi(\zeta)(I_\infty x) = I_0 K(\zeta)x \quad \forall x \in H_0, \quad \zeta \in S. \quad (5)$$

Для характеристической функции $\chi(\zeta)$ справедливы свойства:

1. $K(\zeta)$ — регулярная оператор-функция параметра ζ , $|\zeta| \neq 1$;
2. $\|K(\zeta)\| \leq 1$, если $|\zeta| < 1$ и $\|K(\frac{1}{\zeta})\| < |\frac{1}{\zeta}|$, если $|\zeta| > 1$;
3. $K(\zeta) = K^*(\frac{1}{\zeta}) \quad \forall \zeta, |\zeta| \neq 1$.

Свойства 1 и 2 непосредственно следуют из свойств характеристической функции $\chi(\zeta)$ согласно формуле (5).

Теперь перепишем соотношение (3) в виде

$$\begin{aligned} & [\{I_\infty, I_\zeta\} - \{\Phi^*(\zeta)I_\infty, I_\zeta\}]C(U, \zeta) = \\ & = \{\Phi(\zeta)I_\infty, \bar{R}_\zeta^0 U\} - \{I_\infty, \bar{U}_{\Phi(\zeta)} \bar{R}_\zeta^0 U\}. \end{aligned}$$

Выполним преобразования

$$\begin{aligned} & (\bar{U}_{\Phi(\zeta)} \bar{R}_\zeta^0 U)x = U_{\Phi(\zeta)} R_\zeta(Ux) = \\ & = \frac{1}{\zeta} [R_\zeta^0(Ux) - Ux + V_\zeta C(U, \zeta)x] = \\ & = \frac{1}{\zeta} [(\bar{R}_\zeta^0 - \bar{I})U + V_\zeta C(U, \zeta)]x \quad (x \in H). \end{aligned}$$

Далее имеем

$$\begin{aligned} & \{I_\infty, \frac{1}{\zeta} [(\bar{R}_\zeta^0 - \bar{I})U + V_\zeta C(U, \zeta)]\} = \\ & = \{\Phi^*(\zeta)I_\infty, \bar{R}_\zeta^0 + V_\zeta C(U, \zeta)\} \end{aligned}$$

или

$$\begin{aligned} & [\{\Phi^*(\zeta)I_\infty, V_\zeta\} - \frac{1}{\zeta} \{I_\infty, V_\zeta\}]C(U, \zeta) = \\ & = \frac{1}{\zeta} \{I_\infty, (\bar{R}_\zeta^0 - I)U\} - \{\Phi^*(\zeta)I_\infty, \bar{R}_{\zeta\eta}^0 U\}. \end{aligned}$$

В силу равенства (4) последнее соотношение перепишем в виде

$$\begin{aligned} & [\{I_\infty, V_\zeta\} - \zeta w(\zeta)\{I_0, V_\zeta\}]C(U, \zeta) = \\ & = \zeta[w(\zeta) - w^0(\zeta)]I_0^* \bar{R}_\zeta^0 U \quad (|\zeta| < 1). \end{aligned}$$

С учетом ранее введенных обозначений полученное равенство принимает вид

$$\begin{aligned} & [\Gamma(\infty, \zeta) - \zeta w(\zeta)\Gamma(0, \zeta)]C(U, \zeta) = \\ & = \zeta[w(\zeta) - w^0(\zeta)]I_0^* \bar{R}_\zeta^0 U. \end{aligned}$$

Принимая во внимание характеристическую функцию $K(\zeta)$ оператора V , это соотношение можно записать в таком виде:

$$\begin{aligned} & [E - w(\zeta)K(\zeta)]\Gamma(\infty, \zeta)C(U, \zeta) = \\ & = \zeta[w(\zeta) - w^0(\zeta)]I_0^* \bar{R}_\zeta^0 U. \end{aligned} \quad (6)$$

Так как $\|K(\zeta)\| < 1$ и $\|w(\zeta)\| \leq 1$, то $\forall \zeta \in S$ существует ограниченный обратный оператор $[E - w(\zeta)K(\zeta)]^{-1}$, заданный на всем H . $[E - w(\zeta)K(\zeta)]^{-1}$ является регулярной в S псевдоскалярной функцией от параметра ζ . Отметим, что $\forall \zeta \in S$

$$\operatorname{Re} [E - w(\zeta)K(\zeta)]^{-1} > 0. \quad (7)$$

Действительно, положим

$$M = [E - w(\zeta)K(\zeta)]^{-1}.$$

Так как $\|K(\zeta)\| < 1$ и $\|w(\zeta)\| \leq 1$, $\forall \zeta \in S$, то $\operatorname{Re} M^{-1} = \operatorname{Re}[E - w(\zeta)K(\zeta)] > 0$. Следовательно, из равенства

$$\begin{aligned} \operatorname{Re} M &= \frac{1}{2}[M + M^*] = \frac{1}{2}M(M^{*-1} + M^{-1})M^* = \\ &= \frac{1}{2}M \operatorname{Re} M^{-1} M^* \end{aligned}$$

следует (7).

Теперь из равенства (5) находим

$$\begin{aligned} C(U, \zeta) &= \zeta \Gamma(0, \zeta)^{-1} [E - W(\zeta)K(\zeta)]^{-1} \times \\ &\times [w(\zeta) - w^0(\zeta)]I_0^* \bar{R}_\zeta^0 U. \end{aligned}$$

Далее подставляя в формулу (1) окончательно получим

$$\bar{R}_\zeta U = \bar{R}_\zeta^0 + \zeta \Gamma(0, \zeta)^{-1} [E - W(\zeta)K(\zeta)]^{-1} \times \quad (8)$$

$$\times [W(\zeta) - W^0(\zeta)] I_0^* \bar{R}_\zeta^0 U, \quad (|\zeta| < 1),$$

где W — произвольная оператор-функция класса $K(s; H, H)$.

Итак, резюмируем окончательный результат.

Теорема 2. Формула (8) устанавливает биективное соответствие между совокупностью всех обобщенных резольвент R_ζ оператора V и совокупностью всех псевдоскалярных функций $w(\zeta)$ класса $K(s; H, H)$.

§3. Об одном классе обобщенных резольвент изометрического оператора

Пусть замкнутый изометрический оператор V в гильбертовом пространстве H не имеет ненулевых неподвижных элементов.

Рассмотрим класс обобщенных резольвент R_ζ ($|\zeta| \neq 1$) оператора V , соответствующих унитарным расширениям $V \subset U$ без нулевых неподвижных элементов.

Обозначим через X оператор из N_0 в N_∞ , заданный следующим образом

$$X\varphi = \psi, \quad \text{если } \varphi \in N_0, \psi \in N_\infty \text{ и } \varphi - \psi \in D_V$$

(см. [7,15]). Оператор X является изометрическим и называется запретным относительно оператора V .

Оператору X соответствует изометрический оператор C_x в H_0 такой, что

$$X(V_0 y) = V_\infty(C_x y), \quad \forall y \in D_{C_x}.$$

Оператор C_x тоже будем называть запретным относительно V .

Положим:

$$S_\varepsilon = \left\{ \zeta \in S : -\frac{\pi}{2} + \varepsilon < \arg(1 - \zeta) < \frac{\pi}{2} - \varepsilon, 0 < \varepsilon < \frac{\pi}{2} \right\}.$$

Напомним [7,15], что оператор-функция $\Phi(\zeta)$ класса $K(s; N_0, N_\infty)$ называется допустимой относительно V , если из условий

$$\lim_{\zeta \rightarrow 1, s_\varepsilon} \Phi(\zeta)\varphi = X\varphi$$

$$\text{и } \lim_{\zeta \rightarrow 1, s_\varepsilon} \left[\left| \frac{1 + \zeta}{1 - \zeta} \right| (\|\varphi\| - \|\varphi(\zeta)\varphi\|) \right] < \infty$$

следует, что $\varphi = 0$.

Как показано в [15], совокупность обобщенных резольвент R_ζ ($|\zeta| \neq 1$) оператора V , соответствующих унитарным расширениям $V \subset U$ без нулевых неподвижных элементов, описывается совокупностью всех допустимых относительно V функций $\Phi(\zeta)$ класса $K(s; N_0, N_\infty)$.

Аналогично функцию $W(\zeta)$ класса $K(s; H_0, H_0)$ назовем допустимой относительно V , если из условий

$$\lim_{\zeta \rightarrow 1, s_+} W(\zeta)\varphi = X\varphi$$

и

$$\lim_{\zeta \rightarrow 1, s_+} \left[\left| \frac{1+\zeta}{1-\zeta} \right| (\|\varphi\| - \|W(\zeta)\varphi\|) \right] < \infty$$

следует, что $\varphi = 0$.

Далее отметим, что формулой

$$\Phi(\zeta)I_0x = I_\infty w(\zeta)x, \quad x \in H_0$$

устанавливается биективное соответствие между совокупностью всех допустимых относительно V функций $\Phi(\zeta)$ класса $K(s; N_0, N_\infty)$ и совокупностью всех допустимых относительно V функций $W(\zeta)$ класса $K(s; H_0, H_0)$.

С учетом вышеуказанных заключений приведем итоговый результат.

Теорема 3. *Формула*

$$\begin{aligned} \bar{R}_\zeta U &= \bar{R}_\zeta^0 + \zeta \Gamma(0, \zeta)^{-1} [E - W(\zeta)K(\zeta)]^{-1} \times \\ &\times [W(\zeta) - W^0(\zeta)] I_0^* \bar{R}_\zeta^0 U, \quad (|\zeta| < 1) \end{aligned}$$

устанавливает биективное соответствие между совокупностью всех обобщенных резольвент $R_\zeta (\zeta \in S)$ оператора V , соответствующих унитарным расширениям $V \subset U$ без нулевых неподвижных элементов и совокупностью всех допустимых относительно V функций $W(\zeta)$ класса $K(s; H_0, H_0)$.

В частности, полученная формула позволяет описать решения $F(t)$ ($0 \leq t \leq 2\pi$) операторной тригонометрической проблемы моментов Кратеодори, не имеющих скачка в точке $t = 0$.

Литература

- [1] Крейн М.Г. Об эрмитовых операторах с дефект-индексами, равными единице. — ДАН СССР, 43, N8, 1944. С.339-342.
- [2] Наймарк М.А. О спектральных функциях симметрического оператора. — Известия АН СССР, сер. матем., 7, 1943. С.285-296.
- [3] Крейн М.Г. О резольвентах эрмитова оператора с индексом дефекта (m, m) . — ДАН СССР, 52, N8, 1946. С.657-660.
- [4] Штраус А.В. Обобщенные резольвенты симметрических операторов. — Изв. АН СССР, сер. матем., 18, N1, 1954. С.51-86.
- [5] Штраус А.В. О расширениях, характеристических функциях и обобщенных резольвентах симметрических операторов. — ДАН СССР, 178, N4, 1968. С.790-792.

- [6] Саакян Ш.Н. К теории резольвент симметрического оператора с бесконечными дефектными числами. – ДАН АН Арм. ССР, 41, №4, 1955. С.193-198.
- [7] Штраус А.В. Расширения и обобщенные резольвенты неплотно заданного симметрического оператора. – Изв. АН СССР, сер. матем., 34, N1, 1970. С.175-202.
- [8] Чумакин М.Е. Об обобщенных резольвентах изометрического оператора. – ДАН СССР, 154, N4, 1964. С.254-260.
- [9] Березанский Ю.М. Разложение по собственным функциям самосопряженных операторов. – Наукова думка, Киев, 1965.
- [10] Ильмушкин Г.М. Об обобщенных резольвентах симметрических операторов. – Волжский матем. сб., 13, 1972. С.16-27.
- [11] Александров Е.Л., Ильмушкин Г.М. Обобщенные резольвенты симметрических операторов с произвольными дефектными числами. – Матем. зам. АН СССР, т.19, вып.5, 1976. С.783-794.
- [12] Александров Е.Л., Ильмушкин Г.М. Обобщенные резольвенты симметрических и изометрических операторов. – Известия ВУЗов, сер. матем., N1, Казань, 1977. С.14-23.
- [13] Ильмушкин Г.М., Александров Е.Л. О решениях операторной степенной проблемы моментов. – Известия ВУЗов, сер. матем., N3, Казань, 1989. С.18-23.
- [14] Маркелов Ю.С. К теории характеристических функций изометрических операторов. – Известия ВУЗов, N3, 1967. С.66-73.
- [15] Чумакин М.Е. Об обобщенных резольвентах изометрического оператора, порожденных унитарными расширениями без ненулевых неподвижных элементов. – Тезисы докладов 8-ой научно-технической конф., посвященной 50-летию образования СССР, Ульяновск, 1972. С.242-244.

ЗАДАНИЕ КУСОЧНО-НЕПРЕРЫВНЫХ ФУНКЦИЙ И ДЕЙСТВИЯ НАД НИМИ

П.А. Новиков

Аналитические выражения кусочно-непрерывных функций должны быть удобны для дифференцирования и интегрирования. Для решения этой задачи предлагаются способы ломаных линий и парабол, "сопутствующих" множителей и асимптот.

1. **Способ ломаных линий и парабол.** Ломанная линия, составленная из прямолинейных отрезков (рис. 1), выражается формулой

$$y = k_0 x + a_0 + \sum_{1 \leq i \leq n} k_i |x - x_i| \quad (1)$$

где x_i — абсцисса точки перелома;
 k_i — угловой коэффициент, определяющий величину излома в точке i . Угловые коэффициенты определяются по формулам

$$k_0 = \frac{k_{Л} + k_{П}}{2}; k_1 = \frac{k_{12} - k_{Л}}{2}; \dots$$

$$k_i = \frac{k_{i,i+1} - k_{i-1,i}}{2}; \dots k_n = \frac{k_{П} - k_{n-1,n}}{2}$$

где $k_{i,i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$.

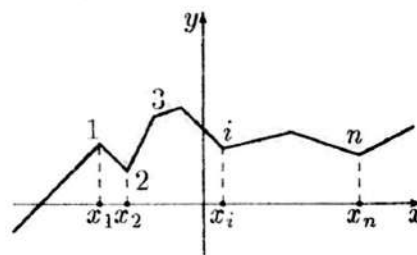


рис.1

Каждый элемент $k_i |x - x_i|$ легко проинтегрировать и дифференцировать воспользовавшись равенством $|x - x_i| = \sqrt{(x - x_i)^2}$:

$$\int |x - x_i| dx = \int \sqrt{(x - x_i)^2} dx = \frac{1}{2} \sqrt{(x - x_i)^2} (x - x_i) = \frac{1}{2} |x - x_i| (x - x_i).$$

В процессе интегрирования каждый прямолинейный отрезок превращается в отрезок квадратной параболы, а точка перелома $x = x_i$ в точку сопряжения. При последующем интегрировании

$$\int \frac{1}{2} |x - x_i| (x - x_i) dx = \frac{1}{6} |x - x_i| (x - x_i)^2$$

квадратная парабола превращается в кубическую, а точка сопряжения — в точку сопряжения второго порядка.

Производная выражения (1) — ступенчатая функция

$$E(x) = k_0 + \sum_{1 \leq i \leq n} k_i \operatorname{sgn}(x - x_i). \quad (2)$$

В каждой точке i функция $E(x)$ имеет скачок на величину $2k_i$.

Предположим, что ломанная линия в точке перелома i имеет дополнительный скачок на величину h_i . Тогда выражение (1) приобретет вид

$$y = k_0 x + a_0 + \sum_{1 \leq i \leq n} k_i \operatorname{sgn}(x - x_i) \left(x - x_i + \frac{h_i}{2k_i} \right). \quad (1')$$

Интегрирование (1') следует производить по формуле 2 из раздела 4.

2. Способ сопутствующих множителей. Аналитическое выражение кусочно-непрерывной функции $f(x)$ на отрезке $a < x < b$ получается ее умножением на "кусочную единицу"

$$e(a, b) = \frac{1}{2} [\operatorname{sgn}(x - a) - \operatorname{sgn}(x - b)].$$

На этой основе можно получить многокусочную функцию

$$y = \sum_{1 \leq i \leq n} \frac{1}{2} [\operatorname{sgn}(x - a_i) - \operatorname{sgn}(x - b_i)] \cdot f_i(x)$$

Пусть $\int f_i(x) dx = F_i(x)$, тогда

$$\int y dx = \sum_{1 \leq i \leq n} \frac{1}{2} [\operatorname{sgn}(x - a_i) (F_i(x) - F_i(a_i)) - \operatorname{sgn}(x - b_i) (F_i(x) - F_i(b_i))].$$

Изменяя сопутствующие множители можно конструировать различные кусочно-непрерывные функции. Например, запишем функцию $|x - a|f(x)$ в виде $\operatorname{sgn}(x - a) \cdot [(x - a)f(x)]$ и определяя $F(x) = \int (x - a)f(x) dx$ найдем

$$\int |x - a|f(x) dx = \int \operatorname{sgn}(x - a) [(x - a)f(x)] dx = \operatorname{sgn}(x - a) [F(x) - F(a)]$$

3. Способ асимптот. Пусть асимптоты задаются двумя произвольно расположенными прямыми f_1 и f_2 . Уравнение гиперболы, привязанной к этим линиям, будет выражаться формулой

$$y = \frac{1}{2} (f_1 + f_2 \pm \sqrt{(f_1 - f_2)^2 + \Delta})$$

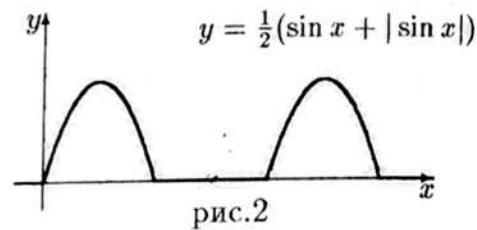
Параметр Δ характеризует степень приближения гиперболы к асимптотам. При $\Delta = 0$ гипербола сливается со своими асимптотами. Верхняя

$$y_{12} = \frac{1}{2} (f_1 + f_2 + \sqrt{(f_1 - f_2)^2}) = \frac{1}{2} (f_1 + f_2 + |f_1 - f_2|). \quad (3)$$

Соответственно, нижняя ветвь

$$y_{21} = \frac{1}{2} (f_1 + f_2 - |f_1 - f_2|). \quad (4)$$

В принципе, в качестве асимптотических линий могут быть приняты любые кривые линии. Для стыкуемых кусков таких линий остаются справедливыми формулы (3) и (4). Например $f_1 = \sin x$ и $f_2 = 0$, тогда верхней ветвью, согласно (3), будет функция изображенная на рис.2.



Способ асимптот позволяет конструировать различные кусочно - непрерывные функции, в том числе и пространственные. Если в качестве f_1 и f_2 будут представлены две пересекающиеся асимптотические поверхности, то формулы (3) и (4) будут выражать собой сплайны, составленные из кусков пространственных поверхностей.

4. Интегрирование кусочно-непрерывных функций. Рекомендуются следующие формулы интегралов для кусочно-непрерывных функций:

1. $\int |x| x^n dx = \frac{|x| x^{n+1}}{n+2} + C, n \neq -2$
2. $\int \operatorname{sgn}(x+a)(x+b) dx = |x+a| \left(\frac{x}{2} - \frac{a}{2} + b \right)$
3. $\int |x+a|(x+b) dx = |x+a|(x+a) \left(\frac{x}{3} - \frac{a}{6} + \frac{b}{2} \right)$
4. $\int \operatorname{sgn}(x-a)f(x) dx = \operatorname{sgn}(x-a)[F(x) - F(a)],$ где $F'(x) = f(x)$
5. $\int |\ln x| dx = \operatorname{sgn}(\ln x)x(\ln x - 1) + \operatorname{sgn}(x-1)$
6. $\int |\sin x| dx = -|\sin x| \operatorname{ctg} x + \sum_{k \in \mathbb{Z}} [\operatorname{sgn}(x - \pi k) + \operatorname{sgn}(k)]$
7. $\int |\cos x| dx = |\cos x| \operatorname{tg} x + \sum_{k \in \mathbb{Z}} [\operatorname{sgn}(x - \frac{\pi}{2}(2k+1)) + \operatorname{sgn}(k)]$

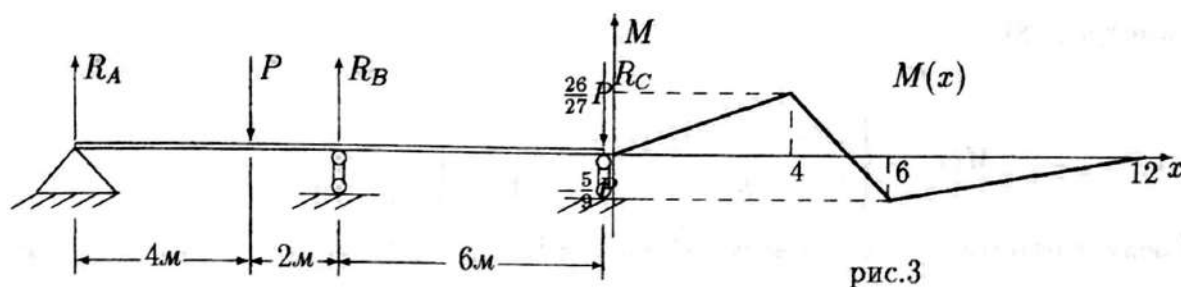
С введением в употребление кусочно-непрерывных функций формула неопределенного интегрирования примет вид :

$$\int f(x) dx = F(x) + E_f(x) + C,$$

где $E_f(x)$ - зависящая от подынтегральной функции $f(x)$ ступенчатая функция, которая призвана скомпенсировать скачки функции $F(x)$ (см. например 4 - 7).

5. Пример использования кусочно-непрерывных функций для решения задач строительной механики. Задана двухпролетная балка (рис. 3). Определить

1. Реакции R_A, R_B, R_C .
2. Изгибающие моменты в точках перелома.



Задача статически неопределимая, но ее можно решить воспользовавшись уравнением прогибов балки (см. [1], [2]):

$$EJy(x) = \iint M(x)dx^2 + Cx + D$$

Решение. Изгибающий момент $M(x)$ в произвольном сечении – кусочно - непрерывная функция вида (1)

$$M(x) = \frac{R_A}{2}|x| - \frac{P}{2}|x-4| + \frac{R_B}{2}|x-6| - \frac{R_C}{2}|x-12| \quad (5)$$

Для интегрирования (5) воспользуемся формулой 1 из раздела 4.

$$\begin{aligned} \iint M(x)dx &= \int \frac{R_A}{4}|x|x - \frac{P}{4}|x-4|(x-4) + \frac{R_B}{4}|x-6|(x-6) - \frac{R_C}{4}|x-12|(x-12)dx \\ &= \frac{R_A}{12}|x|x^2 - \frac{P}{12}|x-4|(x-4)^2 + \frac{R_B}{12}|x-6|(x-6)^2 - \frac{R_C}{12}|x-12|(x-12)^2. \end{aligned}$$

Константы D и C можно найти из граничных условий.

$$EJy(x) = \begin{cases} -\frac{P}{12}4^3 + \frac{R_B}{12}6^3 - \frac{R_C}{12}12^3 + D = 0, & x = 0 \\ \frac{R_A}{12}6^3 - \frac{P}{12}2^3 - \frac{R_C}{12}6^3 + 6C + D = 0, & x = 6 \\ \frac{R_A}{12}12^3 - \frac{P}{12}8^3 + \frac{R_B}{12}6^3 + 12C + D = 0, & x = 12 \end{cases}$$

Отсюда следствие:

$$-54R_A + \frac{70}{3}P - 18R_B + 54R_C = 0. \quad (6)$$

Обратимся снова к выражению (5).

$$M(x) = \begin{cases} -2P + 3R_B - 6R_C = 0, & x = 0 \\ 6R_A - 4P + 3R_B = 0, & x = 12 \end{cases} \quad (7)$$

Решая совместно уравнения (6), (7), находим значения реакций опор :

$$R_A = \frac{13}{54}P; \quad R_B = \frac{23}{27}P; \quad R_C = \frac{5}{54}P. \quad (8)$$

Изгибающий момент в произвольной точке определяется по формуле (5), с учетом параметров (8).

$$M(x) = \begin{cases} 2R_A + R_B - 4R_C, & x = 4 \\ 3R_A - R_B - 3R_C, & x = 6 \end{cases} \stackrel{(8)}{=} \begin{cases} \frac{26}{27}P, & x = 4 \\ -\frac{5}{9}P, & x = 6 \end{cases}$$

Следует отметить, что для всех $x < 0$ и $x > 12$ $M(x) = 0$ (см. (5)).

ЛИТЕРАТУРА

- [1] В.И. Федосьев Соппротивление материалов, -М.: "Наука", 1970.
- [2] А.В. Дарков, Н.Н. Шапошников Строительная механика, -М.: "Высшая Школа", 1986.

СИСТЕМА МУЛЬТИАГЕНТНОЙ ОПТИМИЗАЦИИ ДЛЯ ЗАДАЧИ КОММИВОЯЖЕРА

Н.В. Романов

Введение

Система мультиагентной оптимизации (*Multi-agent optimization system* или *Ant – System*, в дальнейшем *AS*) – это недавно разработанные методы решения задач комбинаторной оптимизации, которые были успешно применены в частности к задаче коммивояжера (в дальнейшем *ЗК*). К сожалению, данное направление не нашло достаточно сильного развития в нашей стране, по крайней мере мною не было встречено ни одной работы, посвященной изучению *AS*. Сразу стоит отметить, что имеющиеся на данный момент методы не имеют сильной математической базы и построены в основном на эвристических соображениях. Тем не менее, изучение данного подхода представляет значительный интерес, так как структура алгоритмов поражает своей гибкостью и возможностью комбинации с точными методами решения.

Модель и структура *AS* была подсказана поведением муравьев в колониях, в частности их поведением при питании (отсюда и название – *Ant – System*). Одной из основных идей является идея ненаправленной коммуникации между агентами (искусственными муравьями), основанная на следах феромона (феромон используется насекомыми, в частности муравьями, для коммуникации). Следы феромона – это своеобразный тип числовой информации, используемый насекомыми для сохранения полученного опыта при решении частных задач.

Первые алгоритмы *AS* [7,8] были применены к *ЗК*, однако с того времени были предложены значительные изменения базовых алгоритмов. В большинстве случаев наиболее быстрые алгоритмы представляют собой комбинацию базовых алгоритмов с алгоритмами локального поиска, улучшающими сгенерированные решения. В данной работе будет дан краткий обзор *AS*, а так же будут рассмотрены различные модификации базовых алгоритмов. В параграфе 2 дается краткая формулировка задачи коммивояжера. Параграфы 3 – 7 посвящены описанию самих алгоритмов, параграф 8 посвящен обобщению полученных результатов.

Задача коммивояжера

$ЗК$ – интенсивно изучается на протяжении последних лет многими направлениями математики. Нужно отметить, что решение данной задачи потребовало значительных исследовательских усилий, однако до сих пор не существует *точного и быстрого* метода решения задач большой размерности. $ЗК$ так же играет важную роль в $АС$, так как первые алгоритмы $АС$ [7,8], а так же их последующие модификации [9,6,16,17] первоначально были применены именно к $ЗК$. Данная задача была выбрана по следующим причинам:

- это задача к которой $АС$ легко применима
- это NP – сложная задача
- это стандартная тестовая модель для новых алгоритмов и хорошие результаты для $ЗК$ обычно рассматриваются как доказательство их *производительности*
- задача легко воспринимаема, что не требует реализации сложных технических моментов

За краткостью содержания я не буду останавливаться на точных определениях теории графов, которые обычно используются для формулировки данной задачи. Интуитивно $ЗК$ – это задача коммивояжера, который хочет начиная со своего родного города, найти самый короткий путь через множество городов и вернуться в свой город. Более формально $ЗК$ может быть представлена связным взвешенным графом $G = (N, A)$, где N – множество вершин, соответствующее городам, а A – множество рёбер, связывающих вершины графа G . Каждому ребру $a_{ij} \in A$ сопоставлено число $d_{ij} \in R$, которое является *длиной* ребра $a_{ij} \in A$ или расстоянием между городами i и j , $i, j \in N$. $ЗК$ – это задача нахождения Гамильтонова цикла минимальной длины, где Гамильтонов цикл – это замкнутый цикл проходящий через каждую вершину $n = |N|$ графа G только один раз.

Если расстояние между городами не зависит от направления движения, т.е. $d_{ij} \neq d_{ji}$, то мы имеем симметричную $ЗК$, в противном случае мы имеем асимметричную или несимметричную $ЗК$.

Описание принципа работы базового алгоритма $АС$ В базовом алгоритме $АС$ (в дальнейшем $АС_{base}$), в случае $ЗК$, построение решения задачи производится путём перемещения агентов из города в город на проблемном графе. На построение решений оказывает влияние *искусственный след феромона*, а так же эвристическая информация, полученная на предыдущих итерациях алгоритма. При применении $АС_{base}$ к $ЗК$, каждому ребру a_{ij} графа G ставится в соответствие определённое количество феромона $\tau_{ij}(t)$, где τ_{ij} – это числовая информация, модифицируемая во

время работы алгоритма, а t – номер итерации. При применении AS_{base} к симметричной $3K$, $\tau_{ij}(t) = \tau_{ji}(t)$, в случае несимметричной задачи $\tau_{ij}(t) \neq \tau_{ji}(t)$

На начальной стадии алгоритма, каждый из m агентов помещается в произвольно выбранный город, а затем каждый агент применяет правило изменения своего местонахождения. Все агенты строят туры следующим образом. Находясь в городе i , агент выбирает ещё не посещённый город j основываясь на величине феромона $\tau_{ij}(t)$, соответствующего ребру a_{ij} и локально доступной эвристической информации, которая является функцией от ребра a_{ij} . Агенты выбирают с большей вероятностью города, которые расположены *ближе* к текущему и соединены рёбрами с большим количеством феромона. Для получения допустимого решения каждому агенту выделяется ограниченная область памяти, называемая *табу листом*, в которой записан текущий тур. *Табу лист* используется для определения непосещённых городов, а так же даёт возможность заново пройти тур и сделать необходимые изменения.

После того, как все агенты построили туры, происходит обновление феромона. Обычно, сначала выполняется понижение феромона на фиксированную величину, а затем каждый агент изменяет количество феромона на посещённых рёбрах. Обновление происходит таким образом, что рёбра содержащиеся в наиболее коротких турах и посещённые большим количеством агентов, получают большее количество феромона, что в свою очередь повышает вероятность их выбора на следующих итерациях алгоритма. В данном контексте количество феромона $\tau_{ij}(t)$ представляет собой изменяемую (и можно даже сказать *обучаемую*) вероятность выбора города j , при нахождении агента в городе i .

Алгоритмы AS_{base} , дающие наилучшие результаты, обычно *улучшают* полученные решения с помощью методов локального поиска.¹ В общем, модель работы этих алгоритмов может быть представлена как последовательное *улучшение* произвольно сгенерированного начального решения. Алгоритмы прекрасно работают для задач небольшой размерности, однако довольно часто, при применении к задачам большой размерности, полученные результаты далеки от оптимальных. Это связано с тем, что при произвольной генерации, полученные начальные решения, вообще говоря, могут и не *привести* алгоритм к туру, хотя бы *близкому* к оптимальному. Обычно эта проблема решается путём значительного увеличения количества итераций, что приводит к увеличению затрат машинного времени. [6,15,17], применяя последний к турам, построенным агентами. Такая комбинация является очень полезной, так как алгоритмы локального поиска очень часто дают плохие решения [12]. Дело в том, что при применении методов локального поиска, использующих произвольно сгенерированные начальные решения, полученный результат может очень сильно отличаться от оптимального [12]. Итак, с одной стороны, применение методов локального по-

¹К методам локального поиска относятся такие широко известные методы как *2-opt*, *3-opt*, *tabu search*, *Lin-Kernighan algorithm* и т.д.

иска может улучшить решение, построенное AS_{base} , а с другой стороны алгоритмы AS_{base} дают более перспективные начальные решения для методов локального поиска. Начальные решения сгенерированные методами AS_{base} более перспективны ввиду того, что искусственными агентами используются те рёбра, которые с большей вероятностью будут включены в оптимальные² туры.

На Рис1. представлена схема работы алгоритма AS_{base} . После инициализации

procedure AS_{base} algorithm for TSP

Set parameters, initialize pheromone trails

while (termination condition not met) **do**

ConstructSolutions

ApplyLocalSearch %optional

UpdateTrails

end

end AS_{base} algorithm for TSP

Рис 1. Схема работы алгоритма AS_{base} для ЗК.

следа феромона и других параметров повторяется основной цикл, до достижения условия останова алгоритма. В основном цикле агенты сначала строят допустимые решения, затем эти решения улучшаются методами локального поиска и в конце цикла идёт обновление феромона. В основном, все алгоритмы AS_{base} дающие наилучшие результаты следуют этой схеме [6,11,14,15,17].

Алгоритм AS_{cycle}

Первоначально было предложено три различных модификации AS_{base} [3,7] – AS_{cycle} , $AS_{quantity}$ и $AS_{density}$. В $AS_{quantity}$ и $AS_{density}$ агенты производят обновление феромона сразу после перемещения из города в город, тогда как в AS_{cycle} изменение следа феромона происходит после построения всех туров, причём количество вкладываемого каждым агентом феромона является значением некоторой функции от длины тура. Необходимо отметить, что экспериментальные данные, полученные при тестировании алгоритма AS_{cycle} , оказались намного лучше данных для алгоритмов $AS_{quantity}$ и $AS_{density}$, поэтому мы остановимся на рассмотрении только этого метода. В AS_{cycle} каждый из m искусственных агентов строит решения по нижеприведённому алгоритму.³

Построение тура. Первоначально, каждый агент помещается на произвольно

²Здесь и далее под *оптимальным* туром будет пониматься один из лучших туров, тогда как самый лучший тур будет назван *строгим оптимальным*. Это связано с тем, что алгоритм не всегда достигает строго оптимального тура.

³Описываемый алгоритм не содержит методов локального поиска, хотя их добавление и приведет к заметному усилению метода.

выбранный город. На каждой итерации алгоритма агент k применяет правило выбора следующего города. Вероятность того, что агент k на шаге t , находящийся в городе i выберет город j задаётся формулой

$$p_{ij}^k = \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in N_i^k} [\tau_{il}(t)]^\alpha [\eta_{il}]^\beta} \quad \text{если} \quad j \in N_i^k \quad (1)$$

Здесь $n_{ij} = 1/d_{ij}$ – это эвристическая информация, полученная на предыдущих итерациях алгоритма, α и β два параметра, определяющие взаимное соотношение феромона и эвристической информации, а N_i^k – это множество городов, *допустимых* для агента k данным шагом.⁴ Параметры α и β имеют следующее значение; если $\alpha = 0$ то большее предпочтение отдаётся городам, находящимся ближе к текущему, если $\beta = 0$, то алгоритм работает используя только информацию о феромоне. Однако при данном значении параметра β алгоритм быстро приходит к стагнации (или остановке) с соответствующей генерацией туров, которые на самом деле сильно *субоптимальны*. Стагнация поиска в [8] определена как процесс, при котором все агенты следуют по одному и тому же пути, что приводит к построению *одинаковых* решений. Итак, здесь существует взаимное влияние феромона и эвристической информации.

Обновление феромона. После построения всех туров, происходит обновление следа феромона. Первоначально производится уменьшение феромона всех рёбер на постоянную величину, затем каждый агент добавляет количество феромона, определённое формулой (2), на все рёбра, содержащиеся в его туре.

$$\tau_{ij}(t+1) = (1-p)\tau_{ij}(t) + \sum_{k=1}^m \Delta\tau_{ij}^k(t) \quad (2)$$

Здесь $0 < p \leq 1$ – это параметр определяющий величину *испарения* феромона. Параметр p используется для ограничения аккумуляции большого количества феромона на некоторых участках, а так же позволяет алгоритму забыть сгенерированные ранее *плотные* решения. Если ребро не содержится ни в одном туре, количество феромона, соответствующее этому ребру, уменьшается экспоненциально. $\Delta\tau_{ij}^k(t)$ – это количество феромона, *вкладываемого* k – м агентом на ребро a_{ij} , содержащееся в его туре. $\Delta\tau_{ij}^k(t)$ определяется следующим образом

⁴При применении методов локального поиска множество N_i^k может и не совпадать с множеством *допустимых* городов. В частности, оно может быть сформировано в результате работы дополнительного алгоритма и являться лишь некоторым подмножеством всех *допустимых* вершин графа G для данного шага. При применении методов локального поиска происходит как бы *улучшение* множества N_i^k , т.е., в это множество включаются города, которые с большей вероятностью будут принадлежать оптимальному туру.

$$\Delta\tau_{ij}^k(t) = \begin{cases} 1/L^k(t) & \text{если } a_{ij} \text{ используется агентом } k \\ 0 & \text{в противном случае.} \end{cases} \quad (3)$$

Здесь $L^k(t)$ – это длина тура k -ого агента. Исходя из равенства (3), чем оптимальней полученный агентом тур, тем больше феромона будет внесено агентом на рёбра, принадлежащие этому туру. Таким образом, рёбра которые используются большим количеством агентов и содержащиеся в наилучших турах получают большее количество феромона, что увеличивает вероятность их выбора на последующих итерациях.

В [8] было произведено сравнение экспериментальных результатов работы алгоритма AS_{cycle} , полученных для задач небольшой размерности (содержащих не более 75 городов), с результатами других эвристических методов. Несмотря на довольно неплохие начальные результаты, для задач большой размерности AS_{cycle} давал довольно плохие решения. Первое усиление алгоритма, названное авторами *элитной стратегией* AS_{cycle} , было предложено в [8,9]. Идея этого алгоритма заключается в дополнительном увеличении феромона рёбер, принадлежащих лучшему туру найденному с момента $t = 0$. Этот тур был обозначен как T^{gb} (*global – best*). Увеличение производилось путём добавления к феромону рёбер, принадлежащих T^{gb} , числа ϵ/L^{gb} , где ϵ – число *элитных* агентов. Некоторые краткие результаты, представленные в [9] демонстрируют, что применение *элитной стратегии* позволяет алгоритму AS_{cycle}

- находить более оптимальные решения
- находить эти решения за меньшее количество итераций.

Однако в этих работах так же представлены результаты, показывающие, что при большом количестве *элитных* агентов, пространство поиска намного раньше сужается к субоптимальным решениям, что приводит к стагнации процесса.

Алгоритм AS_{colony}

Алгоритм AS_{colony} впервые был представлен в [6,10] как непосредственное улучшение AS_{cycle} . AS_{colony} базируется на ранее предложенном теми же авторами алгоритме $AS - Q$, являющемся комбинацией AS_{cycle} с Q – обучением, специфичном типе *самообучающихся* алгоритмов.

AS_{colony} отличается от AS_{cycle} тремя важными аспектами

- AS_{colony} использует более *агрессивное* правило выбора
- феромон *добавляется* только принадлежащим глобально лучшему туру рёбрам
- каждый раз, когда агент использует ребро a_{ij} для перемещения из города i в город j , он уменьшает количество феромона, соответствующее ребру a_{ij} .

Построение тура. В AS_{colony} агенты при выборе следующего города используют *псевдовероятностное* правило выбора. Находясь в городе i , агент k перемещается с вероятностью q_0 в город l , для которого значение $[\tau_{il}(t)]^\alpha [\eta_{il}]^\beta$ является максимальным. Таким образом, наилучшее возможное перемещение определяется по равенству (1) с вероятностью $(1 - q_0)$, которая зависит от полученных ранее следа феромона и эвристической информации.

Глобальное обновление феромона. В AS_{colony} только *глобально* лучший агент изменяет количество феромона после каждой итерации. Это производится в соответствии с равенством (4), которое является модификацией (2).

$$\tau_{ij}(t+1) = (1-p)\tau_{ij}(t) + p\Delta\tau_{ij}^{gb}(t) \quad (4)$$

Здесь $\Delta\tau_{ij}^{gb}(t) = 1/L^{gb}$. Важно обратить внимание на то, что в отличие от AS_{cycle} обновление феромона происходит только на рёбрах, принадлежащих *глобально – лучшему* туру. Параметр p , как и в равенстве (2), представляет *испарение* феромона. Первоначально для обновления следа феромона применялась идея использования лучшего решения каждой итерации. Несмотря на то, что для задач небольшой размерности разница в качестве решений была незначительна, при применении алгоритмов к задачам большой размерности метод использующий *глобально – лучшие* туры даёт более качественные решения.

Локальное обновление феромона. В дополнение к *глобальному* правилу изменения феромона в AS_{colony} так же используется *локальное* правило изменения феромона, которое применяется непосредственно сразу после прохождения агентом ребра во время построения тура,

$$\tau_{ij} = (1-\xi)\tau_{ij} + \xi\tau_0 \quad (5)$$

где $\xi, 0 < \xi < 1$ и τ_0 два параметра. Цель данного изменения заключается в *понижении* вероятности выбора данного ребра другими агентами, что увеличивает вероятность исследования ещё не использованных рёбер.

Единственное различие между AS_{colony} и $AS - Q$ заключается в выборе параметра τ_0 . $AS - Q$ использует формулу вычисления τ_0 , которая выведена из теории Q – обучения. Однако экспериментально было установлено, что замещение τ_0 на константу приводит лишь к незначительному ухудшению решений, что привело к более широкому использованию AS_{colony} в связи со сложностью программной реализации алгоритма $AS - Q$.

Минимаксный алгоритм AS

Минимаксная AS (в дальнейшем $AS_{Max-Min}$) – это прямое улучшение AS_{cycle} . Построение решений в $AS_{Max-Min}$ производится таким же образом как и в AS_{cycle} , т.е. в соответствии с равенством (1), однако в $AS_{Max-Min}$ используется *псевдовероятностное* правило выбора алгоритма AS_{colony} . При использовании данного правила

выбора, довольно хорошие решения могут быть получены очень быстро, но качество достигнутого оптимального тура будет хуже. Основная модификация представленная в $AS_{Max-Min}$ заключается в следующем

- для усиления лучших найденных решений, после каждой итерации только один агент(как и в AS_{colony}) производит обновление феромона
- для избежания стагнации поиска допустимое количество феромона ограничено интервалом $[\tau_{min}, \tau_{max}]$, т.е. $\forall \tau_{ij}, \tau_{min} \leq \tau_{ij} \leq \tau_{max}$
- след феромона при $t = 0$ приравнивается максимальному значению, что приводит к более сильному испарению на начальных стадиях алгоритма [15].

Обновление феромона. После построения туров всеми агентами, след феромона обновляется в соответствии с равенством,

$$\tau_{ij}(t+1) = (1-p)\tau_{ij}(t) + \Delta\tau_{ij}^{best} \quad (6)$$

где $\tau_{ij}^{best} = 1/L^{best}$. Агент, изменяющий количество феромона, может быть выбран как по лучшему итерационному туру T^{ib} (iteration – best), так и по глобально – лучшему туру T^{gb} . В данном алгоритме, чем в большем числе оптимальных туров содержится ребро, тем большее количество феромона оно получает после обновления. Экспериментальные результаты показали, что более оптимальные решения алгоритм даёт при увеличении частоты использования T^{gb} .

Ограничение количества феромона. В $AS_{Max-Min}$ нижняя и верхняя границы введены для избежания стагнации поиска. Введение границ приводит к ограничению вероятности p_{ij} выбора города j во время нахождения агента в городе i интервалом $[p_{min}, p_{max}]$, где $0 < p_{min} \leq p_{ij} \leq p_{max} \leq 1$. $p_{min} = p_{max}$ только в том случае, когда агент имеет только один допустимый город для изменения своего местонахождения. Экспериментальные результаты [15] показали, что нижняя граница следа феромона играет более важную роль чем верхняя, так как максимальное значение феромона ограничено в силу его испарения (что особенно важно при длительном процессе поиска, так как неограниченная аккумуляция феромона приводит к стагнации).

Ранговая AS

Другое улучшение AS_{cycle} – ранговая система AS(в дальнейшем AS_{rank}). В AS_{rank} , для обновления феромона, как и в $AS_{Max-Min}$ используются глобально – лучший тур. Однако помимо этого к обновлению феромона допускается фиксированное количество лучших агентов. Для этого агенты сортируются по длине тура ($L^1(t) \leq L^2(t) \leq \dots \leq L^m(t)$), а количество феромона, которое может вложить каждый агент, вычисляется в соответствии с его рангом r . Только $(w-1)$ лучших агентов каждой итерации изменяют след феромона. Глобально – лучшему агенту

соответствует вес w , r – i -ый агент на каждой итерации производит обновление в соответствии с его весом, заданным как $\max\{0, w - r\}$. Таким образом, изменение феромона происходит по следующему равенству,

$$\tau_{ij} = (1 - p)\tau_{ij}(t) + \sum_{r=1}^{w-1} (w - r)\Delta\tau_{ij}^r(t) + w\Delta\tau_{ij}^{gb}(t) \quad (7)$$

где $\Delta\tau_{ij}^r(t) = 1/L^r(t)$, а $\Delta\tau_{ij}^{gb}(t) = 1/L^{gb}$. В [6] экспериментальные результаты работы алгоритма AS_{rank} , были сравнены с результатами AS_{cycle} , элитной стратегией AS_{cycle} и генетическими алгоритмами. Для задач большой размерности (самая большая задача размерности 250) AS_{rank} и элитная стратегия AS_{cycle} давали лучшие результаты, чем генетические алгоритмы и AS_{cycle} . Необходимо отметить, что результаты полученные AS_{rank} лучше чем результаты полученные элитной стратегией AS_{cycle} , хотя разница незначительна.

Обобщение полученных результатов

Детальное изучение AS показывает, что система имеет большой потенциал решения NP – сложных задач комбинаторной оптимизации. Описанные выше алгоритмы в основном представляют усиленные методы AS , которые дают сильно превосходящие по качеству решения. Основной идеей этих алгоритмов является использование оптимальных туров для последующих итераций. Это достигается сопоставлением большего веса лучшим турам для обновления следа феромона, а так же возможностью вкладывать дополнительное число феромона на рёбра, принадлежащие глобально – лучшему туру. Например в элитной стратегии AS_{cycle} и в AS_{rank} глобально – лучший тур оказывает значительное влияние на обновление феромона (в AS_{rank} в обновлении феромона участвует так же некоторое число лучших агентов); в AS_{colony} и $AS_{Max-Min}$ только один агент, или глобально – лучший или лучший на итерации, участвует в изменении феромона. Более частое использование оптимальных туров даёт лучшие результаты, так как рёбра содержащиеся в этих турах, получают дополнительное число феромона, что увеличивает вероятность их выбора на следующих итерациях.

Однако при использовании усиленных методов существенную роль играет вопрос стагнации процесса. Под стагнацией понимается процесс движения всех агентов по одному и тому же пути или использования одних и тех же рёбер при построении туров. Данная проблема решается путём наложения дополнительных требований на обновление феромона. В AS_{colony} применяется правило локального обновления феромона, суть которого заключается в понижении уровня феромона соответствующего рёбрам, что делает рёбра менее привлекательными для других агентов. В $AS_{Max-Min}$ стагнация поиска предотвращается путём ограничения интервала изменения феромона, соответствующего рёбрам. Здесь вероятность выбора ребра не может быть ниже некоторой фиксированной величины.

Все предложенные модификации AS_{base} , для $ЗК$ дают сильно превосходящие по качеству решения [8,9]. Так же все эти алгоритмы используют оптимальные решения, найденные агентами, на последующих итерациях. Всё это приводит к мысли, что концентрация поиска на пространстве оптимальных решений может дать довольно хорошие результаты.

Описание пространства поиска $ЗК$ может дать ответ на этот вопрос. В работах [2,1,13] были получены результаты описывающие взаимосвязь между длиной тура и расстоянием от строго оптимального тура (расстояние от строго оптимального тура определяется как число различающихся в двух турах рёбер). На Рис2. представлены экспериментальные данные, полученные алгоритмом $AS_{Max-Min}$. Эксперимент проводился для 1000 случайно сгенерированных экземпляров *Евклидовой ЗК*, содержащей 500 городов. Значениями по оси x является расстояние от строго оптимального тура, значением по оси y является длина тура. Из графика видно, что существует взаимосвязь между длиной тура и расстоянием от строго оптимального тура, т.е. чем ближе оптимальное решение к строго оптимальному, тем соответственно меньше длина тура.

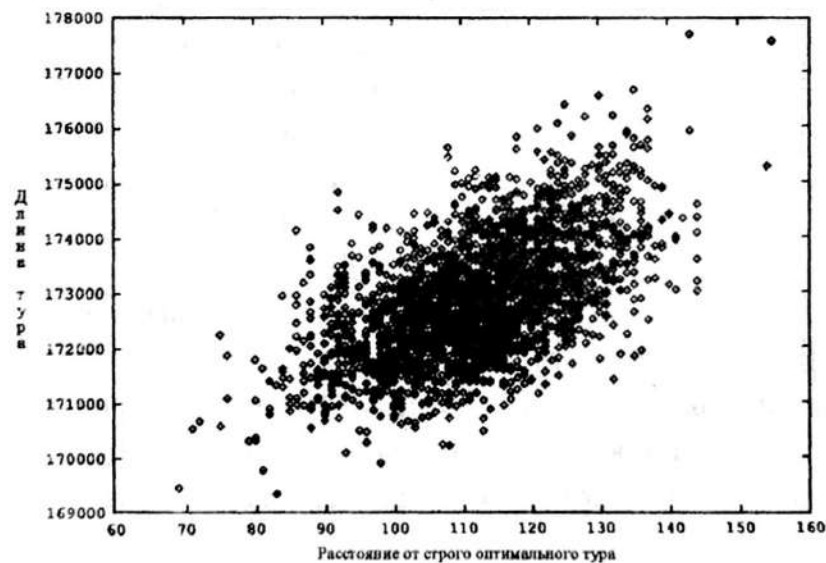


Рис 2. График зависимости между длиной тура и расстоянием до строго оптимального тура. Экспериментальные данные получены для произвольно сгенерированных экземпляров *Евклидовой ЗК* размерности 500.

Задачей алгоритмов AS является генерация перспективных начальных решений, что приводит к существенному сокращению пространства поиска. Интуитивно понятно, что чем лучше начальное решение, тем больше вероятность найти лучшее решение, находящееся *рядом* с начальным. Представленные результаты показывают, что оптимальные решения сосредоточены в некоторой области, с соответствующим коэффициентом корреляции. Поэтому использование данного пространства точ-

ными методами поиска является весьма перспективным направлением.

Литература

- [1] *K.D. Boese* Models for Iterative Global Optimization. PhD thesis, University of California, Computer Science Department, Los Angeles, 1996.
- [2] *K.D. Boese, A.B. Kahng, and S. Muddu* A New Adaptive Multi – Start Technique for Combinatorial Global Optimization. *Operations Research Letters*, 16: pp. 101 – 113, 1994.
- [3] *A. Colorni, M. Dorigo, and V. Maniezzo* Distributed Optimization by Ant Colonies. In *Proceedings of the First European Conference on Artificial Life (ECAL 91)*, pp. 134 – 142. Elsevier, 1991.
- [4] *G. Di Caro and M. Dorigo* AntNet: Distributed Stigmergetic Control for Communications Networks. *Journal of Artificial Intelligence Research (JAIR)*, 9: pp. 317 – 365, December 1998. Available at <http://www.jair.org>.
- [5] *H. Bersini, M. Dorigo, S. Langerman, G. Seront, and L. Gambardella* Results of the First International Contest on Evolutionary Optimisation. In *Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC'96)*, pp. 611-615. IEEE Press, 1996.
- [6] *M. Dorigo and L.M. Gambardella* Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem. *IEEE Transactions on Evolutionary Computation*, 1(1): pp. 53 – 66, 1997.
- [7] *M. Dorigo, V. Maniezzo, and A. Colorni* Positive Feedback as a Search Strategy. Technical Report 91 – 016, Dip. Electronica, Politecnico di Milano, 1991.
- [8] *M. Dorigo, V. Maniezzo, and A. Colorni* The Ant System: Optimization by a Colony of Cooperating Agents. *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, 26(1): pp. 29 – 42, 1996.
- [9] *L.M. Gambardella and M. Dorigo* Ant – Q: A Reinforcement Learning Approach to the Traveling Salesman Problem. In *Proceedings of the Eleventh International Conference on Machine Learning*, pp. 252 – 260. Morgan Kaufmann, 1995.
- [10] *L.M. Gambardella and M. Dorigo* Solving Symmetric and Asymmetric TSPs by Ant Colonies In *Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC'96)*, pp. 622 – 627. IEEE Press, 1996.

- [11] *L.M. Gambardella and M. Dorigo* HAS-SOP: Hybrid Ant System for the Sequential Ordering Problem. Technical Report IDSIA 11 – 97, IDSIA, Lugano, Switzerland, 1997.
- [12] *D.S. Johnson and L.A. McGeoch* The Travelling Salesman Problem: A Case Study in Local Optimization. In E.H.L. Aarts and J.K. Lenstra, editors, *Local Search in Combinatorial Optimization*, pp. 215 – 310. John Wiley, 1997.
- [13] *T. Jones and S. Forrest* Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms. In L.J. Eshelman, editor, *Proceedings of the 6th International Conference on Genetic Algorithms*, pp. 184 – 192. Morgan Kaufman, 1995.
- [14] *V. Maniezzo* Exact and Approximate Nondeterministic Tree – Search Procedures for the Quadratic Assignment Problem. Technical Report CSR 98-1, Scienze dell'Informazione, Università di Bologna, Sede di Cesena, 1998.
- [15] *T. Stützle* Local Search Algorithms for Combinatorial Problems - Analysis, Improvements, and New Applications. PhD thesis, Darmstadt University of Technology, Department of Computer Science, 1998.
- [16] *T. Stützle and H.H. Hoos* Improvements on the Ant System: Introducing the MAX-MIN Ant System. In R.F. Albrecht G.D. Smith, N.C. Steele, editor, *Artificial Neural Networks and Genetic Algorithms*, pp. 245 – 249. Springer Verlag, Wien New York, 1998.
- [17] *T. Stützle and H.H. Hoos* MAX – MIM Ant System and Local Search for Combinatorial Optimization Problems. In S. Voss, S. Martello, I.H. Osman, and C. Roucairol, editors, *Meta – Heuristics: Advances and Trends in Local Search Paradigms for Optimization*, pages 313 – 329. Kluwer, Boston, 1999.

Полилинейные тождества матричных супералгебр типа $M_k \otimes G_u$ над полем характеристики нуль.

Ю. А. Тимаков

Введение.

В современной алгебре большой интерес представляет изучение объектов, называемых супералгебрами – алгебр, объединяющих в себе как коммутирующие, так и антикоммутирующие компоненты. Можно отметить, что многие их свойства оказались полезными и в теоретической физике – идеи суперсимметрии нашли широкое применение в современной теории поля[4]. Одной из важнейших алгебраических задач в исследовании свойств супералгебр является отыскание идеала их тождеств. Для многих объектов данного типа это достаточно трудная задача, поскольку объемы вычислений, связанные с их исследованиями очень велики. Однако на данный момент большое развитие получает применение мощной вычислительной техники, и многие вопросы, носившие до сих пор гипотетический характер, могут в настоящее время быть разрешены при помощи ЭВМ.

Статья посвящена свойствам тождеств матричных супералгебр конечного и бесконечного ранга над полем характеристики нуль. В первой части работы представлены некоторые тождества матричных супералгебр конечного ранга и обзор общих свойств PI -алгебр. Во второй – представлен результат, полученный при помощи компьютера, дающий точную оценку минимальной степени тождеств алгебры $M_2 \otimes G$.

Тождества в матричных супералгебрах конечного ранга над полем характеристики нуль.

Базовые понятия:

F — фиксированное поле характеристики нуль.

Определение. Пусть R - коммутативное кольцо с единицей. Ассоциативной алгеброй над R называется правый унитарный R - модуль A с дистрибутивной ассоциативной операцией умножения и с единицей.

G_u — алгебра Грассмана ранга $u \in N$ над F . $G_u = A_0 \oplus A_u$, где $A_0 = F$ — тело супералгебры; A_u — нильпотентная часть — ассоциативная алгебра с множеством образующих $X = \{e_1, e_2, \dots, e_u\}$. Операция умножения в G_u задана свойством: $e_i e_j = -e_j e_i$. Базисом G_u являются: единица из F и все слова вида $e_{i_1} e_{i_2} \dots e_{i_k}$, где $k = 1 \dots u$, $1 \leq i_1 < \dots < i_k \leq u$. Обозначим множество базисных элементов G_u как $Basis\ G_u$.

G_∞ (или G) — алгебра Грассмана бесконечного ранга — с бесконечным числом образующих.

G можно представить также в виде $G = G^+ \oplus G^-$, где G^+ — четная часть алгебры Грассмана — подалгебра G , базисом которой являются единица и все слова четной длины. G^+ является центром G . G^- — нечетная часть — подпространство G , базисом которого являются все слова нечетной длины.

M_k — алгебра квадратных матриц порядка k над F . $\text{Basis } M_k = \{e_{ij} \mid i, j = 1 \dots k\}$ (матричные единицы).

Матричной супералгеброй в данной работе называется тензорное произведение $M_k \otimes G_u$ (u в данном случае может быть как натуральным числом, так и символом бесконечности). Важно отметить, что это не единственный пример матричных супералгебр. Базис $M_k \otimes G_u$ — элементы $m \otimes g$, $m \in \text{Basis } M_k$, $g \in \text{Basis } G_u$. Элемент $m \otimes g \in \text{Basis } M_k \otimes G_u$ назовем четным (соответственно нечетным), если g — четный (соответственно нечетный). В дальнейшем весь анализ будет касаться свойств матричных супералгебр $M_k \otimes G_u$ над полем характеристики нуль.

Основные результаты, касающиеся матричных супералгебр конечного ранга.

Предложение 1. Для любых $u \geq 0$, $k \geq 2$ стандартный полилинейный многочлен степени $u+2k$ (S_{u+2k}) является тождеством в $M_k \otimes G_u$.

Предложение 2. Пусть A — произвольная ассоциативная PI-алгебра, $f \equiv 0$ в $A \otimes G_s$, $g \equiv 0$ в $A \otimes G_t$ — полилинейные тождества. Тогда $f \cdot g \equiv 0$ в $A \otimes G_{t+s+1}$.

Предложение 3. Пусть $f \equiv 0$ в $A \otimes G_s$, $g \equiv 0$ в $A \otimes G_t$ полилинейные тождества. Тогда полином $f(x) \mid_{x=g} \equiv 0$ в $A \otimes G_{t+s}$.

Для доказательства данных предложений воспользуемся некоторыми вспомогательными утверждениями.

Некоторые свойства тождеств ассоциативных PI-алгебр над полем характеристики нуль.

Пусть A — ассоциативная PI-алгебра над полем F . I — идеал тождеств алгебры A .

Определение. Пусть P_1 и P_2 — две системы тождественных соотношений. Говорят, что система P_1 равносильна системе P_2 , если в каждой алгебре A , в которой соотношения системы P_1 являются тождественными, соотношения P_2 также тождественны и наоборот.

Определение. Полином свободной ассоциативной алгебры $f(x_1 \dots x_n)$ называется полилинейным, если он линеен по каждой своей переменной.

Теорема. Система тождеств ассоциативной PI-алгебры A над бесконечномерным полем F эквивалентна системе полилинейных тождеств A . [6, стр. 279].

Утверждение 1. Пусть A — ассоциативная PI -алгебра. Полилинейный многочлен f является тождеством в A тогда и только тогда, когда при любой подстановке базисных элементов A его значение равно 0.

Доказательство. Необходимость: Следует из тождественности. Достаточность: пусть $B \subset A$ — базис A . $x^i \in A$. И пусть $\{x_b^i\}_{b \in B}$ — координаты x^i в базисе B . $x^i = \sum_{b \in B} x_b^i \cdot b$. Из

$$\text{полилинейности } f(x^1, \dots, x^n) = \sum_{(b_1, \dots, b_n) \in B^n} x_{b_1}^1 \cdot \dots \cdot x_{b_n}^n \cdot f(b_1, \dots, b_n) = 0 \quad \blacksquare$$

Утверждение 2. Пусть для некоторого n S_n (стандартный полилинейный многочлен) является тождеством в ассоциативной PI -алгебре A над полем F характеристики нуль.

Тогда $\forall t \in \mathbb{N}$ полином $g_t(x_1, \dots, x_n; b) = \sum_{\substack{r_0 + \dots + r_n = t \\ r_i \geq 0}} b^{r_0} \cdot \bar{x}_1 \cdot b^{r_1} \cdot \bar{x}_2 \cdot b^{r_2} \dots \cdot b^{r_{n-1}} \cdot \bar{x}_n \cdot b^{r_n}$ (i) явля-

ется следствием S_n .

Доказательство. Вначале введем вспомогательные термины.

Для произвольных $u \in \mathbb{N}$, $v = 1 \dots 2 \cdot u \cdot n$ определим множество целочисленных выборов $R(u, v) = \{(r_0, \dots, r_n) \in \mathbb{Z}_+^{n+1} \mid r_0, r_n \leq u, r_i \leq 2 \cdot u \ \forall i = 1 \dots n-1, r_0 + \dots + r_n = v\}$.

Для фиксированных $b \in A$, $t \in \mathbb{N}$ определим функцию на F $h(s) = s^0 \cdot b^0 + s^1 \cdot b^1 + \dots + s^t \cdot b^t$.

$$\forall s \in F \quad 0 = S_n(h(s) \cdot x_1 \cdot h(s), \dots, h(s) \cdot x_n \cdot h(s)) = \sum_{p=0}^{2 \cdot tn} s^p \times$$

$$\times \sum_{(r_0, \dots, r_n) \in R(t, p)} b^{r_0} \cdot \bar{x}_1 \cdot b^{r_1} \cdot \bar{x}_2 \cdot \dots \cdot \bar{x}_n \cdot b^{r_n}. \quad \text{char } F = 0 \Rightarrow \forall p = 0 \dots 2 \cdot tn$$

$$g_{t,p}(x_1, \dots, x_n; b) = \sum_{(r_0, \dots, r_n) \in R(t, p)} b^{r_0} \cdot \bar{x}_1 \cdot b^{r_1} \cdot \bar{x}_2 \cdot \dots \cdot \bar{x}_n \cdot b^{r_n} \equiv 0. \Rightarrow$$

$$g_{t,t}(x_1, \dots, x_n; b) = g_t(x_1, \dots, x_n; b) \equiv 0$$

■

Свойства $M_k \otimes G_u$.

Утверждение 3. Пусть A — произвольная ассоциативная PI -алгебра. Тождественность полилинейного многочлена в $A \otimes G_u$ равносильна равенству нулю при любой подстановке базисных элементов из $A \otimes G_\infty$ в которой количество нечетных элементов не превосходит u .

(i) Здесь символ подчеркивания над буквой означает кососимметрический набор соответствующих переменных, символ «+» означает симметричный набор.

Под нулевой степенью ненулевого элемента понимается единица алгебры

Доказательство. Необходимость. Пусть $f(x_1, \dots, x_n) \equiv 0$ — полилинейное тождество в $A \otimes G_u$, $x_1 = m_1 \otimes g_1, \dots, x_n = m_n \otimes g_n$ — базисные элементы из $A \otimes G_u$.

$$\begin{aligned} f(x_1, \dots, x_n) &= \sum_{\sigma \in S_n} f_{\sigma} \cdot x_{\sigma(1)} \cdots x_{\sigma(n)} = \sum_{\sigma \in S_n} f_{\sigma} \cdot m_{\sigma(1)} \cdots m_{\sigma(n)} \otimes g_{\sigma(1)} \cdots g_{\sigma(n)} = \\ &= \sum_{\sigma \in S_n} f_{\sigma} \cdot \varepsilon_{\sigma, \bar{g}} m_{\sigma(1)} \cdots m_{\sigma(n)} \otimes g_1 \cdots g_n = p_{\bar{g}}(m_1, \dots, m_n) \otimes g_1 \cdots g_n = 0 \end{aligned}$$

Где $\varepsilon_{\sigma, \bar{g}}$ — знак перестановки нечетных элементов выборки $\bar{g} = (g_1, \dots, g_n)$. В зависимости от расположения и количества нечетных элементов в выборке $\bar{g}$ выполняется одно из условий: $p_{\bar{g}} \equiv 0$ в A или $g_1 \cdots g_n = 0$. Поскольку ранг супералгебры конечный, максимальное количество нечетных элементов в базисной выборке при условии, что $g_1 \cdots g_n \neq 0$, равно u . Следовательно, для любой базисной выборки $\bar{g}$ из G , в которой количество нечетных элементов не превосходит u , $p_{\bar{g}} \equiv 0$.

Достаточность. Воспользуемся введенными выше обозначениями. Пусть $f(x_1, \dots, x_n) \equiv 0$ при любой подстановке базисных элементов из $A \otimes G_u$, в которой количество нечетных элементов не превосходит u . Тогда для любой базисной выборки $\bar{g}$ из G , в которой количество нечетных элементов не превосходит u , $p_{\bar{g}} \equiv 0$. Поскольку максимальное количество нечетных элементов в базисной выборке G_u при условии, что $g_1 \cdots g_n \neq 0$, равно u , $f \equiv 0$ в $A \otimes G_u$ ■

Следствие. Полилинейный полином $f(x_1, \dots, x_n)$ является тождеством в $A \otimes G_u$ тогда и только тогда, когда он является тождеством в $A \otimes G_n$.

Доказательство основных результатов.

Предложение 1. Для любых $u \geq 0, k \geq 2$ стандартный полилинейный многочлен степени $u+2 \cdot k$ ($S_{u+2 \cdot k}$) является тождеством в $M_k \otimes G_u$

Доказательство. Индукция по u . $u=1$: $S_{2 \cdot k+1}$ является T -следствием $S_{2 \cdot k}$, который является тождеством как в M_k [2] так и в $M_k \otimes G_1$. Пусть $u > 1$ и $\forall j < u$ $S_{j+2 \cdot k}$ является тождеством в $M_k \otimes G_j$. По утверждению 3 достаточно показать равенство нулю при любой подстановке базисных элементов из $M_k \otimes G$, в которой количество нечетных элементов не превосходит u . С другой стороны, по индуктивному предположению, $S_{u+2 \cdot k}$ является следствием тождества $S_{u+2 \cdot k-1} \equiv 0$ в $M_k \otimes G_{u-1}$, следовательно $S_{u+2 \cdot k} \equiv 0$ при любой подстановке базисных элементов, количество нечетных элементов в которой не превосходит $u-1$. Отсюда, для доказательства тождественности $S_{u+2 \cdot k}$ в $M_k \otimes G_u$ достаточно показать равенство нулю при любой подстановке базисных элементов, в которой количество нечетных элементов равно u .

Стандартный полином замкнут относительно любой перестановки переменных, следовательно, не ограничивая общности, можно рассматривать только те подстановки, в которых первые u элементов — нечетные, остальные — четные. $x_i = m_i \otimes g_i$

$$S_{u+2k}(x_1, \dots, x_{u+2k}) = \bar{x}_1 \cdot \dots \cdot \bar{x}_u \cdot \bar{x}_{u+1} \cdot \dots \cdot \bar{x}_{u+2k} - \\ - \bar{x}_1 \cdot \dots \cdot \bar{x}_{u-1} \cdot \bar{x}_{u+1} \cdot \bar{x}_u \cdot \dots \cdot \bar{x}_{u+2k} + \dots + \bar{x}_{u+1} \cdot \dots \cdot \bar{x}_{u+2k} \cdot \bar{x}_1 \cdot \dots \cdot \bar{x}_u = *$$

Поскольку первые u переменных выборки g антикоммутируют, после преобразования выражения кососимметрический набор первых u переменных выборки m преобразуется в симметрический.

$$* = (m_1 \cdot \dots \cdot m_u \cdot \bar{m}_{u+1} \cdot \dots \cdot \bar{m}_{u+2k} - m_1 \cdot \dots \cdot m_{u-1} \cdot \bar{m}_{u+1} \cdot m_u \cdot \dots \cdot \bar{m}_{u+2k} + \\ + \dots + \bar{m}_{u+1} \cdot \dots \cdot \bar{m}_{u+2k} \cdot m_1 \cdot \dots \cdot m_u) \otimes g_1 \cdot \dots \cdot g_{u+2k} = |\text{обозначение}| = \\ = p(m_1, \dots, m_{u+2k}) \otimes g_1 \cdot \dots \cdot g_{u+2k}$$

Полином p является линеаризацией полинома

$$g_i(x_1, \dots, x_{2k}; b) = \sum_{\substack{r_0 + \dots + r_{2k} = i \\ r_i \geq 0}} b^{r_0} \cdot \bar{x}_1 \cdot b^{r_1} \cdot \dots \cdot b^{r_{2k-1}} \cdot \bar{x}_{2k} \cdot b^{r_{2k}}, \text{ который по утверждению 3 яв-}$$

ляется следствием тождества $S_{2,k} \equiv 0$ в M_k ■

Замечание. Известно [1], что стандартный полилинейный многочлен степени $k+1$ является тождеством алгебры размерности k . Размерность $M_k \otimes G_u$ равна $k^2 \cdot 2^u$, то есть для $k=2, u=2$ стандартный полилинейный многочлен степени 17 точно будет тождеством. Предложение 1 позволяет снизить эту довольно грубую оценку, т.к. согласно ему при тех же условиях уже S_6 является тождеством.

Предложение 2. Пусть A — произвольная ассоциативная PI-алгебра, $f \equiv 0$ в $A \otimes G_s$, $g \equiv 0$ в $A \otimes G_t$ — полилинейные тождества. Тогда $f \cdot g \equiv 0$ в $A \otimes G_{t+s+1}$.

Доказательство. При любой подстановке базисных элементов, в которой количество нечетных элементов не превосходит $t+s+1$ в полином $f \cdot g$ возможны два случая:

1: Среди переменных из g окажется не более t нечетных элементов. Тогда поскольку $g \equiv 0$ в $A \otimes G_t$, при этой подстановке $g=0$

2: Среди переменных из g окажется не менее $t+1$ нечетных элементов. Тогда среди переменных из f не более s нечетных элементов. Поскольку $f \equiv 0$ в $A \otimes G_s$, при данной подстановке $f=0$. ■

Пример. $S_4(x_1, \dots, x_4) \cdot S_4(x_5, \dots, x_8)$ является тождеством в $M_2 \otimes G_3$, поскольку полином S_4 является тождеством в $M_2 \otimes G_1$.

Предложение 3. Пусть $f \equiv 0$ в $A \otimes G_s$, $g \equiv 0$ в $A \otimes G_t$ полилинейные тождества. Тогда полином $f(x) |_{x=g} \equiv 0$ в $A \otimes G_{t+s}$.

Доказательство. При любой подстановке базисных элементов, в которой количество нечетных элементов не превосходит $t+s$ в полином $f(x) \mid_{x=g}$ возможны два случая:

- 1: Среди переменных из g окажется не более t нечетных элементов. Тогда поскольку $g=0$ в $A \otimes G_t$, при этой подстановке $g=0$
- 2: Среди переменных из g окажется более t нечетных элементов. Тогда среди переменных из f кроме переменной x окажется не более $s-1$ нечетных элементов. Вместе с переменной- подстановкой $x=g$ их будет не более s . Поскольку $f=0$ в $A \otimes G_s$, при данной подстановке $f=0$. ■

Пример. $S_4(x_1, x_2, x_3, S_4(x_4, \dots, x_7))$ является тождеством в $M_2 \otimes G_2$

Вычисления в $M_2 \otimes G$.

Введение.

В настоящее время стало очень доступным и актуальным использование мощной вычислительной техники в решении многих задач алгебры. Некоторые задачи решить аналитически практически невозможно – приходится пользоваться примитивными рутинными технологиями, суммарная сложность которых превосходит возможности человека. Здесь приходят на помощь вычислительные методы и средства, но подчас объем вычислений настолько велик, что они не под силу даже самым мощным ЭВМ без тщательного предварительного анализа задачи.

Постановка задачи.

Известно, что в $M_2 \otimes G$ над полем характеристики нуль существует тождество степени восемь [Размыслов, 7]. Возможно, это минимальная степень тождеств в $M_2 \otimes G$

Гипотеза-предположение. Минимальная степень тождеств в алгебре $M_2 \otimes G$ над полем характеристики нуль равна восьми.

Для ее доказательства достаточно показать, что над полем нулевой или достаточно большой характеристики в $M_2 \otimes G$ не существует полилинейных тождеств седьмой степени.

Описание вычислений, использованных в доказательстве предположения.

Как доказать, что полилинейных тождеств степени 7 в алгебре $M_2 \otimes G$ над полем характеристики нуль не существует? Воспользуемся следующими свойствами:

- 1) Для произвольной PI- алгебры A тождественность полилинейного полинома степени n в алгебре $A \otimes G$ эквивалентна его тождественности в $A \otimes G_n$
- 2) Полилинейный полином является тождеством тогда и только тогда, когда он равен нулю при любой подстановке базисных элементов.
- 3) Пусть f – полилинейный полином степени n . Если при подстановке некоторой выборки базисных элементов $(x_1=m_1 \otimes g_1, \dots, x_n=m_n \otimes g_n)$ из $A \otimes G$ такой, что

$x_1 \cdot x_2 \cdot \dots \cdot x_n \neq 0$, полином $f(x_1, x_2, \dots, x_n)$ равен нулю, то для любой подстановки u вида $(y_1 = m_1 \otimes g_1^*, \dots, y_n = m_n \otimes g_n^*)$, в которой четность каждой переменной g_i^* соответствует четности соответствующей переменной g_i , $f(y_1, y_2, \dots, y_n)$ также равен нулю.

- 4) Если в алгебре $A \otimes G$ над полем F некоторой ненулевой характеристики нет тождеств степени n , то их нет также при нулевой характеристике поля.

Благодаря этому, критерием тождественности полилинейного многочлена степени n в $M_2 \otimes G$ является равенство нулю при подстановке выборов базисных элементов $M_2 \otimes G$ вида $m \otimes g^0$ длины n , где $m \in \{b_1, b_2, b_3, b_4\}$ — базисный элемент M_2 , g^0 является либо единицей (четный элемент), либо некоторой буквой из $\{e_1, \dots, e_n\}$, причем поскольку равенство полинома нулю не зависит от порядка следования нечетных элементов в выборке, для определенности будем считать, что они упорядочены по возрастанию индексов и индекс каждой последующей буквы, не равной единице, на единицу отличается от индекса предыдущей. Обозначим множество таких выборов из G через VG . Пример: $(1, 1, \dots, 1) \in VG$; $(1, e_1, e_2, 1, \dots, 1, e_3, 1) \in VG$; $(1, e_2, \dots, 1) \notin VG$; $(1, e_1, e_1, \dots) \notin VG$; $(e_1, 1, e_3, \dots) \notin VG$.

Множество базисных выборов длины n из M_2 обозначим через VM . $V = VM \times VG$ — множество выборов из $M_2 \otimes G$.

Каждой выборке из VM можно однозначно сопоставить число. Это осуществляется следующим образом: пусть $(b_{i1}, b_{i2}, \dots, b_{in}) \in VM$. Число MC , соответствующее данной выборке $MC = \sum_{r=1}^n (i_r - 1) \cdot 4^{(r-1)}$. Обратно, пусть $d_0, d_1, \dots, d_{n-1}$ — цифры числа MC в четверичной системе исчисления, тогда для $r=1 \dots n$ $i_r = d_{r-1} + 1$. Отсюда, мощность VM равна 4^n .

Число GC , соответствующее выборке из VG , определяется следующим образом: пусть $(g_1, \dots, g_n) \in VG$; $d_0, \dots, d_{n-1}$ — цифры числа в двоичной записи. Для $r=0 \dots n-1$ $d_r = 0$, если $g_{r+1} = 1$. Если $g_{r+1} = e_{xx}$, $d_r = 1$. Обратно, для $r=1 \dots n$ если $d_{r-1} = 0$, $g_r = 1$; если $d_{r-1} = 1$, $g_r = e_s$, где s — это количество вхождений единицы в двоичной записи числа GC среди первых r цифр. Отсюда, $\#VG = 2^n$, $\#V = 2^{3 \cdot n}$.

Полилинейный многочлен f степени n является тождеством в $M_2 \otimes G$, если он соответствует нетривиальному решению $\{c_\tau \in F \mid \tau \in S_n\}$ однородной системы уравнений

$$(1) \left\{ \sum_{\tau \in S_n} c_\tau \cdot x_{\tau(1)} \cdot \dots \cdot x_{\tau(n)} = 0 \mid (x_1, \dots, x_n) \in V \right\}. \text{ В качестве базиса в } M_2 \text{ возьмем:}$$

$b_1 = e_{11} - e_{22}$; $b_2 = e_{12} + e_{21}$; $b_3 = e_{12} - e_{21}$; $b_4 = e_{11} + e_{22}$. Запишем таблицу умножения, соответствующую ему.

	b_1	b_2	b_3	b_4
b_1	b_4	b_3	b_2	b_1
b_2	$-b_3$	b_4	$-b_1$	b_2
b_3	$-b_2$	b_1	$-b_4$	b_3
b_4	b_1	b_2	b_3	b_4

Строка таблицы соответствует левому сомножителю, столбец – правому. Отсюда видно, что для фиксированной выборки таких базисных элементов значения любых полилинейных мономов могут отличаться лишь знаком. Иными словами, $\forall m = (m_1, \dots, m_n) \in VM \forall \tau \in S_n \ m_{\tau(1)} \cdot \dots \cdot m_{\tau(n)} = \varepsilon_{m,\tau} \cdot m_1 \cdot \dots \cdot m_n, \varepsilon_{m,\tau} = \pm 1$. Аналогичным свойством обладает и множество выборок VG . $\forall g = (g_1, \dots, g_n) \in VG \forall \tau \in S_n \ g_{\tau(1)} \cdot \dots \cdot g_{\tau(n)} = \varepsilon_{g,\tau} \cdot g_1 \cdot \dots \cdot g_n, \varepsilon_{g,\tau} = \pm 1$. Отличие в том, что $\varepsilon_{g,\tau}$ определяется не значением элементов выборки, а расположением ее нечетных элементов.

$$\varepsilon_{g,\tau} = \prod_{1 \leq i < j \leq n} (-1)^{I_{i,j}}, I_{i,j} = I(\tau^{-1}(i) > \tau^{-1}(j)) \cdot I(|g_i| \equiv 1 \pmod{2}) \cdot I(|g_j| \equiv 1 \pmod{2}).$$

Отсюда следует, что система (1) записывается в виде (2)

$$\left\{ \sum_{\tau \in S_n} c_\tau \cdot \varepsilon_{g,\tau} \cdot \varepsilon_{m,\tau} = 0 \mid g \in VG, m \in VM \right\}. \text{ Получилась обычная система линейных од-}$$

нородных уравнений размерности $n! \times \#V$. Для $n=7$ это составляет **5040** столбцов и примерно $2 \cdot 10^6$ строк. Это очень большая размерность и даже отбросив из рассмотрения подстановки, соответствующие заведомо линейно зависимым строкам, уменьшить количество строк системы удалось лишь в три раза. Как показали эксперименты, чтобы перебрать все строки системы, потребовалось бы примерно три недели непрерывной работы компьютера с процессором *Intel Celeron 466*. Однако для доказательства данного предположения достаточно найти первые $7! = 5040$ линейно независимых строк.

Описание алгоритмов.

Поскольку объем вычислений огромен, использовать числа с плавающей точкой нецелесообразно – накопление погрешности вычислений может привести к неверному результату. Вычисления проводились над полем высокой характеристики – над кольцом вычетов Z_{31013} ⁱ⁾.

ⁱ⁾ Такая характеристика выбрана из двух соображений: первое – если в поле некоторой ненулевой характеристики однородная система имеет только тривиальное решение, то и в поле нулевой характеристики она также невырождена; второе – уместительное замечание, что чем больше характеристика поля, тем больше шансов, что невырожденная в нулевой характеристике система будет невырождена над этим

Алгоритм приведения к треугольному виду в кольце вычетов.

Строки системы генерируются по очереди и добавляются в конец треугольной преобразованной матрицы $A=(a_{ij})$. Пусть в матрице A перед добавлением новой строки $s-1$ преобразованных строк. Номер добавляемой строки – s . Алгоритм преобразования строки заключается в последовательном вычитании из нее преобразованных строк с коэффициентами, подобранными так, чтобы индекс лидера преобразуемой строки после вычитания из нее i -й строки оказался больше i . Пусть из строки s вычтена $i-1$ строка. Тогда все ее элементы с индексами меньшими чем i равны нулю. Если $a_{si}=0$, то строка уже является преобразованной; Если $a_{si} \neq 0$, то если индекс лидера i -й строки больше чем i ($a_{ii}=0$), то меняем местами i -ю и s -ю строки, после чего s -я строка становится преобразованной. В случае, если $a_{si} \neq 0$ и $a_{ii} \neq 0$, то умножаем s -ю строку на лидер i -й строки и вычитаем из нее i -ю строку, умноженную на лидер s -й строки $(A)_{s::}=a_{ii} \cdot (A)_s - a_{si} \cdot (A)_i$. (через $(A)_i$ обозначается i -я строка матрицы A). После вычитания лидер s -й строки «обнуляется» и она становится преобразованной.

Если после преобразования строка s станет равной нулю, то ее удаляем – ранг матрицы не увеличился.

Алгоритм применяем дотех пор, когда либо будут перебраны все строки, либо ранг системы строк достигнет количества столбцов.

После применения алгоритма для $M_2 \otimes G$ для седьмой степени в поле Z_{31013} было найдено 5040 ($=7!$) линейно независимых строк, что и доказывает предположение.

Литература

- 1) Ф. Каш «Модули и кольца». Издательство «Мир». Москва 1981.
- 2) Р. Пирс «Ассоциативные алгебры». Издательство «Мир». Москва 1986.
- 3) И. Херстейн «Некоммутативные кольца». Издательство «Мир». Москва 1972.
- 4) Ф. А. Березин «Введение в алгебру и анализ с антикоммутирующими переменными» Издательство МГУ 1983.
- 5) Размыслов Ю.П. «О конечной базирюемости тождеств матричной алгебры второго порядка над полем характеристики нуль». Алгебра и логика. – 1973.- Т.12, № 1.- С. 83-113.
- 6) Мальцев А.А. «Избранные труды. Т1. Классическая алгебра». Москва, Наука 1976.

Спектральное представление вполне неунитарного оператора

Л.А. Штраус

В данной работе получено спектральное представление вполне неунитарного замкнутого изометрического оператора T , действующего в гильбертовом пространстве H . В ней уточняются факты, анонсированные в [1], а также приводятся их доказательства.

Представление линейных операторов в виде операторов умножения на независимую переменную рассматривалось в работах М.Г. Крейна, Б.С. - Нады и Ч. Фояша [2], А.В. Штрауса [3], [4], автора [1], [5] и др. Здесь мы покажем, что методы, развитые в работах А.В. Штрауса, особенно естественно позволяют получить спектральное представление в случае изометрического оператора.

Сохраняя в основном обозначения работ [1], [5], введем необходимые понятия. Для каждого $h \in H$ рассмотрим векторнозначную аналитическую внутри единичного круга функцию

$$\tilde{h}(\zeta) = P_0(I - \zeta T_0^*)^{-1} h; \quad |\zeta| < 1;$$

и линейное пространство $\tilde{H} = \{\tilde{h}(\zeta)\}$ всех таких функций. Здесь T_0 - оператор, совпадающий с T на его области определения D_T и равный нулю на $N = D_T^\perp$; P_ζ , $|\zeta| < 1$, - оператор проектирования в H на $N_0 = R_T^\perp$ параллельно $M_\zeta = (T - \zeta I)D_T$.

Определение. Оператор T в H назовем вполне неунитарным, если в H не существует ненулевого подпространства, на котором T индуцирует унитарный оператор.

Введем в рассмотрение подпространство

$$M = \bigcap_{|\lambda| \neq 1} M_\lambda.$$

Заметим, что имеет место включение $\bigcap_{0 < |\lambda| \neq 1} M_\lambda \subset M_0$, следовательно,

$$M = \bigcap_{0 < |\lambda| \neq 1} M_\lambda. \text{ Более того, } M \subset D_T.$$

Действительно, если $f \in \bigcap_{0 < |\lambda| \neq 1} M_\lambda$, то для любого λ , $|\lambda| \neq 1$, $\lambda \neq 0$ найдется элемент g_λ

такой, что

$$f = (T - \lambda I)g_\lambda.$$

Переходя в последнем равенстве к пределу при $\lambda \rightarrow 0$, получаем $f = Tg_0$. Отсюда следует включение

$$\bigcap_{0 < |\lambda| \neq 1} M_\lambda \subset M_0 \quad (M_0 = R_T).$$

Аналогично, рассматривая обратный оператор T^{-1} вместо T , получаем $M \subset D_T$.

Предложение. Если оператор T вполне неунитарный, то линейное многообразие $\sum_{|\lambda| \neq 1} N_\lambda$ плотно в H , т. е.

$$\overline{\sum_{|\lambda| \neq 1} N_\lambda} = H.$$

Доказательство. Достаточно установить, что если подпространство

$$M = \bigcap_{0 < |\lambda| \neq 1} M_\lambda = \left(\overline{\sum_{|\lambda| \neq 1} N_\lambda} \right)^\perp$$

является нулевым, то T индуцирует на нем унитарный оператор $T|_M$. Унитарность $T|_M$ следует из инвариантности M относительно T и T^{-1} :

если $f \in M_\lambda$, $f = (T - \lambda I)g$, то $Tf = (T - \lambda I)Tg$, т. е. $Tf \in M_\lambda$;

если $f \in M_{T^{-1}, \lambda}$, $f = (T^{-1} - \lambda I)g$, то $T^{-1}f = (T^{-1} - \lambda I)T^{-1}g$, $T^{-1}f \in M_{T^{-1}, \lambda}$; но $M_{T^{-1}, \lambda} = M_{T, 1/\lambda}$, $\lambda \neq 0$, поэтому из принадлежности $f \in M_{T, 1/\lambda}$ следует $T^{-1}f \in M_{T, 1/\lambda}$.

Отображение $\Phi: h \rightarrow \tilde{h}$ пространства H на $\tilde{H}$ в соответствии с терминологией [3] назовем спектральным представлением оператора T .

Из доказанного предложения следует, что для изучения Φ в случае вполне неунитарного оператора достаточно найти образ многообразия $\sum_{|\lambda| \neq 1} N_\lambda$ при спектральном представлении.

Значение характеристической функции $\Theta(\zeta)$ оператора T и оператор P_ζ связаны соотношением

$$\Theta(\zeta) = \zeta P_\zeta |N.$$

Теорема. Образом многообразия $\sum_{|\lambda| \neq 1} N_\lambda$ при представлении Φ является

$$\Phi \sum_{|\lambda| \neq 1} N_\lambda = \sum_{|\lambda| < 1} \frac{I - \zeta \bar{\lambda} P_\zeta (P_\lambda |N)^*}{1 - \zeta \bar{\lambda}} N_0 + \sum_{|\lambda| < 1} \zeta \frac{P_\zeta - P_\lambda}{\zeta - \lambda} N$$

или в терминах характеристической функции Θ

$$\Phi \sum_{|\lambda| \neq 1} N_\lambda = \sum_{|\lambda| < 1} \frac{1 - \Theta(\zeta)\Theta^*(\lambda)}{1 - \zeta\bar{\lambda}} N_0 + \sum_{|\lambda| < 1} \frac{\lambda\Theta(\zeta) - \zeta\Theta(\lambda)}{\zeta - \lambda} N.$$

При этом оператору T отвечает оператор умножения на ζ в $\tilde{H}$.

Доказательство. При $|\lambda| < 1$ подпространство $R_T = M_0$ можно представить в виде

$$R_T = (I - \lambda T_0^*)^{-1} M_\lambda,$$

следовательно,

$$N_\lambda = (I - \bar{\lambda} T_0)^{-1} N_0. \quad (1)$$

Поскольку $P_\lambda = P_0(I - \lambda T_0^*)^{-1}$, то

$$(P_\lambda|N)^* = P_N(I - \bar{\lambda} T_0)^{-1}|N_0 \quad (2)$$

и оператор $(R_\lambda|N)^*$ является сужением на N_0 оператора проектирования в H на N параллельно $M_{1/\lambda}$. Кроме того,

$$(I - \bar{\lambda} T_0)^{-1}|N_0 = I. \quad (3)$$

Из равенств (1) - (3) получаем

$$\begin{aligned} N_\lambda &= (I - \bar{\lambda} T_0)^{-1} N_0 = \\ &= (I - \bar{\lambda} T_0)^{-1} \left((I - (P_\lambda|N)^*) + (P_\lambda|N)^* \right) N_0 = \\ &= \left((I - \bar{\lambda} T)^{-1} (I - (P_\lambda|N)^*) + (P_\lambda|N)^* \right) N_0. \end{aligned} \quad (4)$$

Из равенств

$$\Phi f = \Phi \left((I - \lambda T)(I - \lambda T)^{-1} f \right) = (1 - \lambda \zeta) \Phi (I - \lambda T)^{-1} f$$

следует

$$\Phi (I - \lambda T)^{-1} f = \frac{1}{1 - \lambda \zeta} \Phi f.$$

Отсюда и из (4) получаем

$$\Phi N_\lambda = \left(\frac{1}{1 - \zeta \bar{\lambda}} \Phi (I - (P_\lambda|N)^*) + \Phi (P_\lambda|N)^* \right) N_0 =$$

$$\begin{aligned}
 &= \left(\frac{I - P_{\zeta}(P_{\lambda}|N)}{1 - \zeta\bar{\lambda}} + P_{\zeta}(P_{\lambda}|N)^* \right) N_0 = \\
 &= \frac{I - \zeta\bar{\lambda}P_{\zeta}(P_{\lambda}|N)^*}{1 - \zeta\bar{\lambda}} N_0, \\
 \Phi N_{\lambda} &= \frac{I - \zeta\bar{\lambda}P_{\zeta}(P_{\lambda}|N)^*}{1 - \zeta\bar{\lambda}} N_0. \quad (5)
 \end{aligned}$$

Аналогично, при $0 < |\lambda| < 1$

$$\begin{aligned}
 N_{T, 1/\lambda} &= N_{T^{-1}, \bar{\lambda}} = (I - \lambda T_0^*)^{-1} N = \left((I - \lambda T_0^*)^{-1} (I - P_{\lambda}) + (I - \lambda T_0^*)^{-1} P_{\lambda} \right) N = \\
 &= \left((I - \lambda T^{-1})^{-1} (I - P_{\lambda}) + P_{\lambda} \right) N, \\
 \Phi N_{1/\lambda} &= \left(\frac{\zeta}{\zeta - \lambda} \Phi(I - P_{\lambda}) + P_{\lambda} \right) N = \left(\zeta \frac{P_{\zeta} - P_{\lambda}}{\zeta - \lambda} + P_{\lambda} \right) N.
 \end{aligned}$$

Отсюда и из (5) получаем

$$\Phi \sum_{|\lambda| \neq 1} N_{\lambda} = \sum_{|\lambda| < 1} \frac{I - \zeta\bar{\lambda}P_{\zeta}(P_{\lambda}|N)^*}{1 - \zeta\bar{\lambda}} N_0 + \sum_{|\lambda| < 1} \left(\zeta \frac{P_{\zeta} - P_{\lambda}}{\zeta - \lambda} + P_{\lambda} \right) N. \quad (6)$$

С учетом того, что $P_{\lambda} N \subset N_0$ и подпространство N_0 входит в первое из многообразий в правой части равенства (6), запишем последнее в виде

$$\Phi \sum_{|\lambda| \neq 1} N_{\lambda} = \sum_{|\lambda| < 1} \frac{I - \zeta\bar{\lambda}P_{\zeta}(P_{\lambda}|N)^*}{1 - \zeta\bar{\lambda}} N_0 + \sum_{|\lambda| < 1} \zeta \frac{P_{\zeta} - P_{\lambda}}{\zeta - \lambda} N$$

или, переходя к характеристической функции,

$$\Phi \sum_{|\lambda| \neq 1} N_{\lambda} = \sum_{|\lambda| < 1} \frac{I - \Theta(\zeta)\Theta^*(\lambda)}{1 - \zeta\bar{\lambda}} N_0 + \sum_{|\lambda| < 1} \frac{\lambda\Theta(\zeta) - \zeta\Theta(\lambda)}{\zeta - \lambda} N.$$

Последнее утверждение теоремы следует из того, что

$$\Phi(T - \zeta I)h = P_{\zeta}(T - \zeta I)h = 0,$$

то есть

$$\Phi Th = \zeta \Phi h.$$

ЛИТЕРАТУРА

1. Штраус Л. А. *О спектральном представлении вполне неунитарного оператора*. // *Фундаментальные проблемы математики и механики*. Вып. 2. -Ульяновск, 1996, с. 38-39.
2. Секефальви-Надь Б., Фоояш Ч. *Гармонический анализ операторов в гильбертовом пространстве*. -М.: Мир, 1970.
3. Штраус А. В. *Спектральные представления линейных операторов*. // *Функциональный анализ*. Вып. 34. -Ульяновск, 1993, с. 80-93.
4. Штраус А. В. *Функциональные модели и обобщенные спектральные функции симметрических операторов*. // *Препринт*. -Ульяновск, 1997.
5. Штраус Л. А., Баринова И. В. *К теории представления регулярного оператора*. // *Фундаментальные проблемы математики и механики*. -Ульяновск, 1996, с. 24-32.

ЭФФЕКТИВНЫЕ ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА МОДЕЛИРОВАНИЯ ЛИНЕЙНЫХ СТОХАСТИЧЕСКИХ СИСТЕМ УПРАВЛЕНИЯ

Ю.В. Цыганова, А.Д. Юрьев

Введение.

Задача оптимального управления динамическим объектом, функционирующим в стохастической обстановке, имеет много приложений (измерительные, навигационные системы и т.п.). Дополнительно она характеризуется необходимостью оценивания состояний как разомкнутых систем, так и систем с обратной связью. При этом неотъемлемой частью исследования алгоритмов оптимального управления и оценивания является создание эффективных инструментальных средств для моделирования динамических стохастических систем управления. Такие программные комплексы зачастую заменяют натурные эксперименты, либо компьютерное моделирование практических задач служит предварительным этапом для их проведения.

В данной работе даются краткие характеристики инструментальной среды для моделирования поведения линейных динамических стохастических систем управления. Наряду с ниже перечисляемыми достоинствами важным свойством данной разработки является разделение объекта, измерителя, фильтра и регулятора, что позволяет строить более гибкую схему моделирования и решать достаточно широкий класс задач.

Постановка задачи.

Рассмотрим модель динамической системы в дискретном времени, которая описывается следующими уравнениями состояния объекта и измерителя:

$$x(t_{i+1}) = \Phi(t_{i+1}, t_i)x(t_i) + B_d(t_i)u(t_i) + G_d(t_i)w_d(t_i), \quad i = 0, 1, 2, \dots$$

$$z(t_i) = H(t_i)x(t_i) + v(t_i), \quad i = 1, 2, \dots$$

где $x(t_i)$, $z(t_i)$, $w_d(t_i)$, соответственно, векторы в R^n , R^m , R^s . Шумы $w_d(t_i)$, $v(t_i)$ образуют гауссовы белые последовательности с нулевыми средними и матрицами ковариаций, соответственно, $Q_d(t_i)$, $R(t_i)$. Матрица $\Phi(t_{i+1}, t_i)$ - переходная матрица состояния от момента времени t_i к моменту t_{i+1} . Вектор $x(t_0)$ - начальное значение вектора состояния с математическим ожиданием $\bar{x}_0$ и матрицей ковариации P_0 . Это

так называемая ЛГК задача управления (линейная задача с гауссовыми случайными величинами и квадратичным функционалом качества).

Критерий качества управления:

$$J = \sum_{i=0}^N [x^T(t_i)X(t_i)x(t_i) + u^T(t_i)U(t_i)u(t_i)] + x^T(t_{N+1})X_f x(t_{N+1})$$

$$U > 0, X \geq 0$$

Где U и X соответствующие весовые матрицы регулятора. Необходимо минимизировать данный функционал качества по $u(t_i) \forall i = 0, 1, 2, \dots$

Для данной постановки задачи справедлива теорема разделения.

Теорема (разделения)[1]. В линейных системах с квадратичным функционалом качества и входным гауссовым шумом стохастически оптимальное управляющее устройство представляет собой фильтр Калмана и линейный функциональный преобразователь оценки состояний ($G_r^*(t_i)$) в управляющий параметр.

При этом $G_r^*(t_i)$ подчиняется принципу эквивалентности, т. е. ($G_r^*(t_i)$) в точности соответствует преобразователю в детерминистской задаче управления с полными (точными) измерениями вектора состояния.

Детерминистская задача управления:

$$x(t_{i+1}) = \Phi(t_{i+1}, t_i)x(t_i) + B_d(t_i)u(t_i), \quad i = 0, 1, 2, \dots$$

$$z(t_i) = H(t_i)x(t_i), \quad i = 1, 2, \dots$$

Алгоритм для вычисления $G_r^*(t_i)$ получается с помощью метода динамического программирования:

$$\begin{aligned} A_c(t_i) &= B_d^T(t_i)K_c(t_i)\Phi(t_{i+1}, t_i) \\ G_r^*(t_i) &= [U(t_i) + B_d^T(t_i)K_c(t_i)B_d(t_i)]^{-1} A_c(t_i) \\ K_c(t_i) &= X(t_i) + \Phi^T(t_{i+1}, t_i)K_c(t_i)\Phi(t_{i+1}, t_i) - A_c^T(t_i)G_r^*(t_i) \\ u^*(t_i) &= -G_r^*(t_i)x(t_i) \end{aligned} \quad (1)$$

Начальные значения:

$$K_c(t_{N+1}) = X_f$$

Для $i = N, N-1, N-2, \dots, 1$ производится расчет по формулам (1).

При согласовании данного алгоритма с алгоритмами оптимальной фильтрации действия по моделированию делятся на две части:

- 1) расчет матрицы;
- 2) оптимальная фильтрация с использованием матрицы.

Схема процесса моделирования выглядит следующим образом:

- 1) из объекта в измеритель поступает вектор наблюдений;
- 2) в измерителе по вектору наблюдений рассчитывается оценка вектора состояний $x(t_i)$;
- 3) оценка вектора состояний подается в регулятор в котором происходит вычисление управляющего сигнала $u^*(t_i)$;
- 4) управляющий сигнал подается в объект для вычисления следующего вектора состояний и вектора измерений.

Как видно из схемы процесса получена замкнутая система, для старта которой необходимо знать вектор $u(t_0)$. Этот вектор вводится в процессе формирования (заполнения матриц) модели.

Следует также отметить, что данный алгоритм оптимального управления, в силу того, что на каждом шаге он использует всю информацию о модели и внешних воздействиях, как прошлых так и будущих, является истинным оптимальным управлением.

Таким образом, задачей является разработка программных средств для моделирования и исследования процессов в замкнутых динамических стохастических системах управления.

Возможности интегрированной среды для сравнительного исследования различных вычислительных методов оптимальной дискретной фильтрации и управления.

Была разработана интегрированная среда для изучения и проведения сравнительного анализа различных методов оптимальной дискретной фильтрации, а также метода истинного оптимального управления. Данный программный продукт является развитием ранее разработанного пакета СМОУ[2]. В настоящей разработке произведен учет и доработка выявленных недостатков предыдущей версии пакета, а также добавлены новые возможности. Данный программный продукт имеет следующие свойства.

Интегрированная среда обладает всеми возможностями предыдущей версии, а именно: ввод параметров задачи в дискретной форме любой размерности, сохранение и редактирование данных, проведение расчетов по заданным алгоритмам, сохранение и просмотр результатов расчетов в форме отчетов и графиков.

Пакет обладает удобным пользовательским интерфейсом. Он написан на языке высокого уровня с использованием технологии ООП, обеспечивающей его дальнейшую модификацию и расширение. Главной отличительной чертой данной разработки является то, что она написана с использованием Windows - технологии.

Изменениям подверглись некоторые части пакета: ввод и сохранение данных задачи, сделан упор на облегчение процесса работы пользователя, на наглядность представления данных.

Были в значительной степени изменены средства построения графиков выбранных сигналов по графическим данным, полученным в результате работы одного из реализованных алгоритмов. Кроме того, был расширен сам спектр сохраняемых сигналов.

К внутренним изменениям относится то, что при разработке программы была изменена сама структура внутреннего представления. Следует особо отметить что в данной версии были разделены объекты измерителя, модели, регулятора. Это дает возможность исследовать не только оптимальные алгоритмы фильтрации и управления, но и субоптимальные алгоритмы и алгоритмы обнаружения нарушений и идентификации неизвестных параметров в линейных динамических системах, что является существенным расширением класса моделируемых задач.

Литература

- [1] Саридис Дж. Самоорганизующиеся стохастические системы управления. М.: Наука, 1980.
- [2] Цыганова Ю.В. Разработка программных средств моделирования и исследования стохастических систем оптимальной дискретной фильтрации. // Фундаментальные проблемы математики и механики: Сборник статей. Выпуск 3. Ульяновск: УлГУ, 1997. 133с.

И.В. Семушин, Ю.В. Цыганова, М.В. Куликова

О вычислении функции правдоподобия для гауссовских марковских последовательностей

1. Задачи обнаружения нарушений

Рассмотрим две системы уравнений

$$x_i(k+1) = \Phi_i x_i(k) + \Gamma_i u_i(k), \quad (1.1)$$

$$z_i(k) = H_i x_i(k) + v_i(k), \quad (1.2)$$

одна из которых (при $i = 1$) описывает исходную физическую систему, а другая при $i = 2$ — ту же систему, но с возможными нарушениями, то есть с другими значениями параметров. В уравнениях (1.1) и (1.2) k означает дискретное время, а $x_i(k)$, $z_i(k)$, $u_i(k)$ векторы, соответственно, в R^n , R^m , R^q . Шумы $u_i(k)$ и $v_i(k)$ образуют гауссовы белые последовательности и характеризуются соотношениями

$$E[u_i(k)] \neq 0, \quad E[v_i(k)] = 0, \quad E[u_i(k)v_i^T(j)] = 0,$$

$$E[u_i(k)u_i^T(j)] = Q_i \delta_{kj}, \quad E[v_i(k)v_i^T(j)] = R_i \delta_{kj},$$

где Q_i — неотрицательно определенные матрицы и R_i — положительно определенные матрицы ковариаций, δ_{kj} — символ Кронекера.

Допустим, имеется фильтр

$$\hat{x}(k+1|k) = \Phi_1 \hat{x}(k|k), \quad \hat{x}(k|k) = \hat{x}(k|k-1) + K_1 \nu(k), \quad (1.3)$$

$$z(k) = \begin{cases} z_1(k), & k < k_1, \\ z_2(k), & k \geq k_1, \end{cases} \quad \nu(k) = z(k) - H_1 \hat{x}(k|k-1), \quad (1.4)$$

$$K_i = M_i H_i^T C_i^{-1} \quad (i = 1, 2), \quad C_i = H_i M_i H_i^T + R_i, \quad (1.5)$$

$$P_i = M_i - M_i H_i^T C_i^{-1} H_i M_i, \quad M_i = \Phi_i P_i \Phi_i^T + \Gamma_i Q_i \Gamma_i^T, \quad (1.6)$$

который при $k < k_1$ является установившимся ($K_1 = \text{const}$) и, как видно из (1.5) и (1.6), оптимальным относительно системы (1.1), (1.2) при $i = 1$.

В этих условиях требуется обнаружить, произошло ли переключение с $z_1(k)$ на $z_2(k)$ в (1.4) в момент времени k_1 . В действительности здесь имеются две различные, хотя и тесно связанные задачи [1]. Первая — обнаружение того, от какой системы ($i = 1$ или $i = 2$) поступают данные $z(k)$ в фильтр, начиная с момента k_1 (сам момент k_1 известен). Вторая задача — обнаружение изменения характеристик

системы с первого набора ($i = 1$) на второй набор ($i = 2$) значений (момент k_1 такого нарушения неизвестен). Далее можно рассматривать и более сложную задачу, когда значения характеристик для второго набора неизвестны или могут быть упорядочены как несколько отдельных наборов; тогда определение момента k_1 может трактоваться как контроль момента нарушения оптимальности фильтра [2].

Известно, что для выработки соответствующего правила принятия решения в этих задачах требуется вычисление функции правдоподобия. Целью статьи является формулировка простого алгоритма вычисления этой функции и алгебраическое доказательство соответствующей теоремы.

2. Вычисление логарифмической функции правдоподобия (ЛФП)

Функция правдоподобия для результата i -го наблюдения $z_i = z(t_i)$ в системе (1.1), (1.2) при условии обработки предыдущих результатов наблюдений имеет вид:

$$f(z_i|Z_{i-1}) = [(2\pi)^m \det(C(t_i))]^{1/2} \times \\ \times \exp\left\{-\frac{1}{2}[z_i - H(t_i)\hat{x}(t_i)]^T C^{-1}(t_i)[z_i - H(t_i)\hat{x}(t_i)]\right\},$$

где $C(t_i) = H(t_i)P(t_i^-)H^T(t_i) + I$, а I — единичная матрица. Обозначим обновляющую последовательность через $\nu(t_i) = z_i - H(t_i)\hat{x}(t_i^-)$, которая характеризуется ковариационной матрицей $C(t_i)$. Тогда логарифмическая функция правдоподобия будет иметь вид

$$\lambda = -\frac{m}{2} \ln(2\pi) - \frac{1}{2} \ln\{\det(C(t_i))\} - \frac{1}{2} \nu^T(t_i) C^{-1}(t_i) \nu(t_i).$$

Для эффективного вычисления двух ее слагаемых, а именно, $\ln\{\det(C(t_i))\}$ и $\nu^T(t_i) C^{-1}(t_i) \nu(t_i)$ предлагается следующий алгоритм.

Алгоритм 1:

I. Задать начальные значения:

$$\hat{x}^0(t_i^+) = \hat{x}(t_i^-); \quad P^0(t_i^+) = P(t_i^-);$$

$$\delta^0(t_i) = 0; \quad \Delta^0(t_i) = 0;$$

II. Для $k = 1, 2, \dots, m$ вычислить:

$$\hat{x}^{(k)}(t_i^+) = \hat{x}^{(k-1)}(t_i^+) + K^{(k)}(t_i) \nu^{(k)}(t_i);$$

$$\nu^{(k)}(t_i) = z^{(k)}(t_i) - h_k^T \hat{x}^{(k-1)}(t_i^+);$$

$$K^{(k)}(t_i) = P^{(k-1)}(t_i^+) h_k / \alpha_k;$$

$$\begin{aligned}\alpha_k &= h_k^T P^{(k-1)}(t_i^+) h_k + 1; \\ P^{(k)}(t_i^+) &= P^{(k-1)}(t_i^+) - K^{(k)}(t_i) h_k^T P^{(k-1)}(t_i^+); \\ \delta^{(k)}(t_i) &= \delta^{(k-1)}(t_i) + \frac{[\nu^{(k)}(t_i)]^2}{\alpha_k}; \\ \Delta^{(k)}(t_i) &= \Delta^{(k-1)}(t_i) + \ln(\alpha_k);\end{aligned}$$

III. Выдать результат:

$$\begin{aligned}[\nu^T(t_i) C^{-1}(t_i) \nu(t_i)] &= \delta^{(m)}(t_i); \\ [\ln\{\det(C(t_i))\}] &= \Delta^{(m)}(t_i);\end{aligned}$$

IV. Конец.

Здесь предполагается, что $R = I$. Данное ограничение не является существенным, так как если матрица $R \neq I$, то используются псевдонаблюдения $\bar{z}(t_i) = L^{-1}z(t_i)$, где матрица L получена разложением Холецкого матрицы $R = LL^T$. Для псевдонаблюдений получим уравнение вида (1.2), где для матрицы $H(t_i)$ берется выражение $L^{-1}H(t_i)$.

Теорема 1. Представим матрицу $H(t_i)$ построчно:

$$H(t_i) = \begin{bmatrix} h_1^T \\ h_2^T \\ \dots \\ h_m^T \end{bmatrix}$$

где h_k^T — k -я строка, рассматриваемая как матрица размера $(1 \times n)$, $n = \dim x(t_i)$. Алгоритм 1 обеспечивает вычисление слагаемых $\ln\{\det(C(t_i))\}$ и $\nu^T(t_i)C^{-1}(t_i)\nu(t_i)$ для логарифмической функции правдоподобия, где вектор $z_i = z(t_i)$ обрабатывается поэлементно:

$$z(t_i) = \begin{bmatrix} z(t_i)^{(1)} \\ z(t_i)^{(2)} \\ \dots \\ z(t_i)^{(m)} \end{bmatrix}$$

Доказательство. Введем обозначения: $\hat{x}^{(0)}(t_i^+) = \hat{x}(t_i) = x^0$, $P(t_i^-) = P^{(0)}$. Рассмотрим структуру матрицы $C(t_i)$, которая определяется согласно (1.5) как $C(t_i) = H^T(t_i)P^{(0)}H(t_i) + I$. Имеем

$$C(t_i) = \begin{bmatrix} h_1^T P^{(0)} h_1 + 1 & h_1^T P^{(0)} h_2 & \dots & h_1^T P^{(0)} h_m \\ h_2^T P^{(0)} h_1 & h_2^T P^{(0)} h_2 + 1 & \dots & h_2^T P^{(0)} h_m \\ \dots & \dots & \dots & \dots \\ h_m^T P^{(0)} h_1 & h_m^T P^{(0)} h_2 & \dots & h_m^T P^{(0)} h_m + 1 \end{bmatrix}$$

Часть I. Необходимо доказать справедливость формулы

$$\ln\{\det(C(t_i))\} = \sum_{k=1}^m \ln(\alpha_k), \quad (2.1)$$

где числа α_k , $k = 1, \dots, m$, вычисляются по алгоритму 1.

Доказательство проведем индукцией по размеру матрицы $C(t_i)$.

1. Пусть $k = 2$. Проверим справедливость равенства

$$\ln\{\det(C(t_i))\} = \ln(\alpha_1) + \ln(\alpha_2)$$

или эквивалентно,

$$\det(C(t_i)) = \alpha_1 \cdot \alpha_2,$$

где числа α_i , $i = 1, 2$, вычисляются по алгоритму 1.

$$\alpha_1 \cdot \alpha_2 = c_{11}^{(0)} \cdot (c_{22}^{(0)} - \frac{1}{\alpha_1} [c_{12}^{(0)}]^2) = c_{11}^{(0)} \cdot c_{22}^{(0)} - [c_{12}^{(0)}]^2 = \det(C(t_i)).$$

Здесь $c_{ij}^{(0)}$ обозначает (i, j) -й элемент матрицы $C(t_i)$. Для этого случая часть I теоремы верна.

2. Пусть теорема 1 верна для $k = m$, то есть справедлива формула

$$\det(C(t_i)) = \prod_{j=1}^m \alpha_j, \quad (2.2)$$

где числа α_j , $j = 1, \dots, m$ определяются согласно алгоритму 1.

Проверим справедливость формулы (2.2) для $k = m + 1$. Для этого рассмотрим матрицу $C^{(1)} = H^T P^{(1)} H + I$, где матрица $P^{(1)}$ определяется по алгоритму 1. Имеем

$$C^{(1)} = H^T P^{(0)} H - \frac{1}{\alpha_1} H^T P^{(0)} h_1 h_1^T P^{(0)} H + I,$$

или

$$C^{(1)} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & c_{22}^{(0)} - [c_{12}^{(0)}]^2 / c_{11}^{(0)} & \dots & c_{2,m+1}^{(0)} - c_{21}^{(0)} c_{1,m+1}^{(0)} / c_{11}^{(0)} \\ \dots & \dots & \dots & \dots \\ 0 & c_{m+1,2}^{(0)} - c_{m+1,1}^{(0)} c_{12}^{(0)} / c_{11}^{(0)} & \dots & c_{m+1,m+1}^{(0)} - [c_{m+1,1}^{(0)}]^2 / c_{11}^{(0)} \end{bmatrix}.$$

Видно, что матрица $C^{(1)}$ получена из матрицы $C(t_i)$ методом исключения Гаусса по первому столбцу (или по первой строке в силу симметричности матрицы $C(t_i)$), где ведущая строка (здесь первая) нормирована (здесь ведущий элемент $c_{11}^{(0)}$).

Из структуры матрицы $C^{(1)}$ и свойств детерминанта следует

$$\det(C(t_i)) = \det(C^{(1)}) \cdot c_{11}^{(0)}. \quad (2.3)$$

Обозначим минор матрицы $C^{(1)}$, полученный в результате вычеркивания первой строки и первого столбца, как C^* . Тогда для каждого элемента матрицы C^* справедлива следующая формула:

$$c_{ij}^* = c_{i+1,j+1}^{(0)} - \frac{c_{i+1,1}^{(0)} c_{1,j+1}^{(0)}}{c_{11}^{(0)}}, \quad i, j = 1, \dots, m. \quad (2.4)$$

Для квадратной матрицы C^* размера m выполняется предположение индукции, то есть

$$\det(C^*) = \prod_{i=1}^m \alpha_i^*, \quad (2.5)$$

где числа α_i^* , $i = 1, \dots, m$, вычисляются согласно алгоритму 1, в котором уже сделана одна итерация. Из формул (2.3) и (2.5) с учетом введенных выше обозначений получим окончательный результат

$$\det(C(t_i)) = c_{11}^{(0)} \det(C^*) = \alpha_1 \prod_{i=2}^{m+1} \alpha_i = \prod_{i=1}^{m+1} \alpha_i,$$

где числа α_i , $i = 1, \dots, m+1$, определяются согласно алгоритму 1. Таким образом часть I теоремы доказана.

Часть II. Необходимо доказать справедливость формулы

$$\nu^T(t_i) C^{-1}(t_i) \nu(t_i) = \sum_{i=1}^m \frac{\nu_i^2}{\alpha_i}, \quad (2.8)$$

где левая часть формулы (2.8) характеризует векторную обработку по типу формулы (1.4), а правая часть характеризует покомпонентную (скалярную) обработку по алгоритму 1.

Доказательство опять проведем по индукции.

1. Пусть $k = 2$. Проверим справедливость формулы

$$\nu^T(t_i) C^{-1}(t_i) \nu(t_i) = \sum_{i=1}^2 \frac{\nu_i^2}{\alpha_i},$$

или, что тоже самое

$$\frac{\nu_1^2 A_{11} + 2\nu_1 \nu_2 A_{12} + \nu_2^2 A_{22}}{\det(C(t_i))} = \frac{\nu_1^2}{\alpha_1} + \frac{\nu_2^2}{\alpha_2}, \quad (2.9)$$

где A_{ij} — соответствующее алгебраическое дополнение матрицы $C(t_i)$. Кроме того, здесь использована симметричность матрицы $C(t_i)$. Поскольку левая часть формулы (2.9) характеризует векторную обработку, а правая — скалярную, обозначим через $(\nu_i)_\circ$ — i -ю компоненту вектора $\nu(t_i)$ соответствующую векторной обработке и через $(\nu_i)_{\text{ск}}$ — i -ю компоненту вектора $\nu(t_i)$ соответствующую скалярной обработке. Тогда легко заметить, что

$$(\nu_1)_\circ = (\nu_1)_{\text{ск}} = z^1 - h_1^T x^0,$$

но уже $(\nu_2)_\circ \neq (\nu_2)_{\text{сж}}$, поскольку

$$\begin{aligned} (\nu_2)_\circ &= z^2 - h_2^T x^0 (\pm h_2^T x^1) = z^2 - h_2^T x^1 + h_2^T (x^1 - x^0) = \\ &= (\nu_2)_{\text{сж}} + h_2^T (x^0 + K_1 \nu_1 - x^0) = (\nu_2)_{\text{сж}} + h_2^T K_1 \nu_1. \end{aligned}$$

Таким образом, можно осуществить переход в формуле (2.9) от векторной обработки к скалярной путем подстановки полученных выше результатов в левую часть формулы (2.9). Получим

$$\begin{aligned} &\{(\nu_1^2)_\circ A_{11} + 2(\nu_1)_\circ (\nu_2)_\circ A_{12} + (\nu_2^2)_\circ A_{22}\} / \det(C(t_i)) = \\ &= \{(\nu_1^2)_{\text{сж}} c_{22} - 2(\nu_1)_{\text{сж}} (\nu_2)_{\text{сж}} c_{12} - 2(\nu_1^2)_{\text{сж}} c_{12} h_2^T K_1 + (\nu_2^2)_{\text{сж}} c_{11} + \\ &+ 2(\nu_1)_{\text{сж}} (\nu_2)_{\text{сж}} c_{11} h_2^T K_1 + c_{11} (\nu_1^2)_{\text{сж}} [h_2^T K_1]^2\} / \det(C(t_i)), \end{aligned}$$

где K_1 вычисляется согласно алгоритму 1, то есть $K_1 = P^{(0)} h_1 / \alpha_1$. Примем во внимание, что $\det(C(t_i)) = \alpha_1 \alpha_2$ согласно доказанному в части I утверждению данной теоремы. Тогда имеем

$$\begin{aligned} &\frac{(\nu_2^2)_{\text{сж}} c_{11}}{\det(C(t_i))} - \frac{2(\nu_1)_{\text{сж}} (\nu_2)_{\text{сж}} c_{12}}{\det(C(t_i))} + \frac{2(\nu_1)_{\text{сж}} (\nu_2)_{\text{сж}} c_{11} (h_2^T P^{(0)} h_1)}{\alpha_1 \det(C(t_i))} + \\ &+ \frac{(\nu_1^2)_{\text{сж}}}{\det(C(t_i))} [c_{22} - \frac{2c_{12}^2}{c_{11}} + \frac{c_{12}^2}{c_{11}}] = \frac{(\nu_1^2)_{\text{сж}}}{\alpha_1} + \frac{(\nu_2^2)_{\text{сж}}}{\alpha_2}, \end{aligned}$$

то есть формула (2.9) верна.

2. Пусть теорема 1 верна для $k = m$, то есть для матрицы $C(t_i)$ размера $(m \times m)$ справедлива формула

$$\nu^T(t_i) C^{-1}(t_i) \nu(t_i) = \sum_{j=1}^m \frac{\nu_j^2}{\alpha_j}, \quad (2.10)$$

где числа α_j , $j = 1, \dots, m$, определяются согласно алгоритму 1.

Проверим справедливость формулы (2.10) для $k = m + 1$.

$$\begin{aligned} \nu^T(t_i) C^{-1}(t_i) \nu(t_i) &= \frac{1}{\det C(t_i)} \left[\sum_{j=1}^{m+1} \nu_j \sum_{i=1}^{m+1} \nu_i A_{ij} \right] = \\ &= \frac{1}{\det C(t_i)} \left[\sum_{j=2}^{m+1} \nu_j \sum_{i=2}^{m+1} \nu_i A_{ij} + 2 \sum_{i=2}^{m+1} \nu_i \nu_1 A_{i1} + \nu_1^2 A_{11} \right]. \end{aligned} \quad (2.11)$$

Здесь, как и раньше, A_{ij} обозначает соответствующее алгебраическое дополнение симметричной матрицы $C(t_i)$.

Применим к матрице $C(t_i)$ метод Гаусса исключения неизвестных по первому столбцу и полученную матрицу обозначим через $\hat{C}$. Нетрудно показать, что в этом

случае формула (2.10) остается верна и для матрицы $\hat{C}$. С другой стороны, из формулы (2.11) и с учетом вышеизложенного справедливо

$$\begin{aligned} \nu^T(t_i)\hat{C}^{-1}(t_i)\nu(t_i) &= \frac{1}{\det \hat{C}(t_i)} \left[\sum_{j=2}^{m+1} \nu_j \sum_{i=2}^{m+1} \nu_i \hat{A}_{ij} \right] + \\ &+ \frac{1}{\det \hat{C}(t_i)} \left[2 \sum_{i=2}^{m+1} \nu_i \nu_1 \hat{A}_{i1} + \nu_1^2 \hat{A}_{11} \right], \end{aligned}$$

где $\hat{A}_{ij}$ — соответствующее алгебраическое дополнение матрицы $\hat{C}$. Используя структуру матрицы $\hat{C}$, нетрудно показать, что $\hat{A}_{1j} = 0$, $j = 2, \dots, m+1$. Таким образом,

$$\nu^T(t_i)\hat{C}^{-1}(t_i)\nu(t_i) = \left\{ \sum_{j=2}^{m+1} \nu_j \sum_{i=2}^{m+1} \nu_i \hat{A}_{ij} + \nu_1^2 \hat{A}_{11} \right\} / \det \hat{C}(t_i). \quad (2.12)$$

Тогда с учетом того, что $\hat{A}_{11} = \prod_{j=2}^{m+1} \alpha_j$ (это непосредственно следует из доказательства части I настоящей теоремы) и принятых обозначений, формула (2.12) примет вид

$$\nu^T(t_i)\hat{C}^{-1}(t_i)\nu(t_i) = \frac{\sum_{j=2}^{m+1} \nu_j \sum_{i=2}^{m+1} \nu_i \hat{A}_{ij}}{\det \hat{C}(t_i)} + \frac{\nu_1^2}{\alpha_1}. \quad (2.13)$$

Обозначим минор матрицы $\hat{C}$, полученный в результате вычеркивания первой строки и первого столбца, через матрицу C_* .

Заметим, что $\hat{A}_{ij} = c_{11} A_{i-1, j-1}^*$, $i, j = 2, \dots, m+1$, где A_{ij}^* — соответствующее алгебраическое дополнение матрицы C_* , а значит верно соотношение

$$\sum_{j=2}^{m+1} \nu_j \sum_{i=2}^{m+1} \nu_i \hat{A}_{ij} = c_{11} \det(C_*) \nu_*^T C_*^{-1} \nu_* = \det(\hat{C}) \nu_*^T C_*^{-1} \nu_*,$$

где ν_* — вектор-столбец вида $\nu_*^T = [\nu_2, \nu_3, \dots, \nu_{m+1}]$.

Для матрицы C_* размера $(m \times m)$ выполняется предположение индукции, то есть справедлива формула

$$\nu_*^T C_*^{-1} \nu_* = \sum_{j=1}^m \frac{[\nu_j^*]^2}{\alpha_j^*}, \quad (2.14)$$

где α_i^* , $i = 1, \dots, m$, вычисляются по алгоритму 1, в котором уже сделана одна итерация. С учетом принятых обозначений, из формул (2.13) и (2.14) следует

$$\nu^T(t_i)\hat{C}^{-1}(t_i)\nu(t_i) = \sum_{i=2}^{m+1} \frac{\nu_i^2}{\alpha_i} + \frac{\nu_1^2}{\alpha_1} = \sum_{i=1}^{m+1} \frac{\nu_i^2}{\alpha_i},$$

где α_i , $i = 1, \dots, m+1$, вычисляются согласно алгоритму 1. Теорема доказана.

Список литературы

- [1] Newbold P.M., Ho Yu-Chi. Detection of changes in the characteristics of a Gauss-Markov process// IEEE Trans. Aerosp. and Electr. Syst. 1968. Vol. AES-4. № 5. P. 707-718.
- [2] Семушин И.В. Контроль оптимальности адаптивного фильтра Калмана по реализации скалярного процесса// Техническая кибернетика. 1979. № 6. С. 195-198.

A. Murgu, O. Yu. Gorokhov, I. V. Semoushin

Input-Output Statistical Inference for Switching Processes

Abstract

This paper deals with the on-line change detection problem of Markov Modulated Poisson Processes (MMPPs) in the framework of switching systems. The dependency between the input and output statistical patterns in a black box viewpoint is analyzed. The abrupt change is treated as a transition from a state of the underlying Markov chain to another. We derive a generalized likelihood ratio (GLR) statistical test to detect the abrupt changes, simultaneously with the estimation of unknown distribution parameters using a likelihood scheme. The input-output mapping has a black-box representation in the class of ARMAX models.

Key words. Probabilistic Routing, Markov Modulated Poisson Process, Statistical Inference, Least Squares, Change Detection.

1 Introduction and Motivation

There are two main approaches to finding the relations between input and output of particular system. First one is a model-based approach, in which we should take into account by some reason the model of considered phenomena and describe it with a set of parameters. These parameters are to be estimated via information extraction from statistical experiment. Through this model with estimated parameters one can establish the relation between the input and output of the system. In the black box or statistical approach, we do not assume the inner system functionality be described by some model. Instead of it, models for input-output information flows are considered. The behaviour of the system is represented by all possible changes in the model parameters for output as a result of changing the model parameters for input.

The problem of statistical description of switching stochastic service systems is a topic interesting in itself. The model of input-transactions output of such system is the set of N traffic rate processes which are modelled as Markov modulated Poisson processes (MMPP). An MMPP is a doubly stochastic Poisson process whose intensity rate is modulated by an n -state underlying Markov chain. The MMPP is parameterized by a vector of intensities of the primary Poisson process $\Lambda \in R^n$ and the matrix of transition probabilities $Q \in R^{n \times n}$ of Markov chain. The relation between parameters of input and output MMPP is to be found via jointly performed change detection and parameter estimation.

It is known that CUSUM test gives the infimum of the worst mean delay for a class of statistical test, with given a priori rate of false alarms in the case of known parameters before and after the change. The application of this test to Markov modulated time series is considered in [7]. In the most realistic situation (for example network traffic) the parameters before and after the change are not known. On-line change detection algorithm is based on generalized likelihood ratio (GLR) test for MMPP model of traffic rate process.

2 Problem Statement

Let us define the topology of physical switching network represented as a graph $G_p = (N, A)$, where N is the set of nodes and A is set of links [6]. For each physical link $a \in A$, there are m_a channels with a finite capacity. The logical topology associated with the switching network is represented as a graph $G_l = (N, L)$, where L is the set of logical links. Each logical link $l \in L$ corresponds to a simple path in G_p . Let denote by W the set of origin-destination (O-D) pairs and by P , the set of all logical paths, while P_w is the set of logical paths that use $l = (i, j) \in L$.

The arrival rate of packets for (O-D) pair $w \in W$ at time step t is r_w^t (packets/second), the rate of flow on logical path $p \in P$ x_p^t (packets/second), $t \in [0, T]$, where T is the time horizon. Let us denote by f_{ij}^t (packets/second) the aggregate rate of flows over the logical links connecting the nodes i and j . Each link $a_{ij} = (i, j) \in L$ has a capacity C_{ij}^t at time step $t, t \in [0, T]$.

For all commodities or traffic classes $w \in W$ we assume that input traffic can be represented as n -state MMPP characterized by the set of parameters $\{\Lambda_w^{in}, Q_w^{in}\}_{w \in W}$, $\Lambda_w^{in} \in R^n$, $Q_w^{in} \in R^{n \times n}$. The departure flow traffic is also modelled as n -state MMPP for each commodity with parameters $\{\Lambda_w^{dep}, Q_w^{dep}\}$ for all $w \in W$.

Taking into account capacity constraints of links in network that are reflecting in possible saturation behaviour of departure traffic for less priority commodities, it can be noted that the number of states for model of departure traffic may be reduced. In more complicated cases, some criteria of choosing the number of Markov chain states can be used.

The problem is to analyze both the input and the output traffic patterns determined by the possible changes of the regime that is associated with the intensity of the primary Poisson process. This problem of traffic analysis is considered in the framework of statistical black-box model identification. The problem of on-line detection in traffic regime can be formulated as follows. Let θ_0 - the value of intensity parameter be known before change occurs. Our goal is to determine the change of regime as quick as it is possible (estimating t_a alarm time), second is to find the estimate θ_1 after change. In the continuation of experiment it is assumed that estimated parameter θ_1 on the previous segment of traffic becomes known parameter θ_0 at the current considered segment of traffic till the new switch of regime is determined. In this sequential scheme the error of estimating of parameter θ_1 on the previous segment will not influence on the estimating parameter θ_1 at the current segment.

2.1 Probabilistic Routing and Statistical Modeling

The statistical viewpoint enable us to determine the behaviour of the system working under routing probabilities in each node for a given commodity. At each node, the routing probabilities are computed accordingly to an ordering of forward virtual paths (VPs) at this node, for each pair source-destination [4]. Specifically, let $|N(i, k)| = n_i^k$ be the number of outgoing

links from node i handling traffic for a destination of the commodity k . Thus, we have in each node i a routing table for each commodity. Note that such an ordering provides the a *priority assignment* over the forward VPs (i, j) y , $y = n_i^k$ for the highest priority. Denote the priority of a VP (i, j) for commodity k by π_{ij}^k . Denote by $P_j^k[\text{empty}]$ the conditional probability that a forward VP (i, j) becomes *available* first, for cells of commodity k , when

$$P_j^k[\text{empty}] = \frac{1 - \varrho_{ij}}{|N(i, k)| - \sum_{x=1}^{|N(i, k)|} \varrho_{ix}} \quad (1)$$

where ϱ_{ij} is the utilization of VP (i, j) . It can easily be proved that taking into account that busy periods of the forward VPs are mutually independent, exponentially distributed random variables with parameters $1 - \varrho_{ix}$, $\forall x \in N(i, k)$ and that the probability that VPs (i, j) has the smallest busy period is

$$P_j^k[\varrho_{ij} < \varrho_{ix}] = \frac{1 - \varrho_{ij}}{\sum_{x=1}^{|N(i, k)|} (1 - \varrho_{ix})}, \text{ for all } x \neq j \quad (2)$$

Given the priority assignment and the utilization values of the VPs, we establish the routing probabilities. The routing algorithm will switch a packet belonging to k to the VP (i, j) with priority y if all of the links with higher priority are busy and this VP is available. The probability P_{ij}^k that edge (i, j) is selected for commodity k with priority y is

$$P_{ij}^k = (1 - \varrho_{ij}) \prod_{x: \pi_{ix}^k > y} \varrho_{ix} + \left(\prod_{x=1}^{n_i^k} \varrho_{ix} \right) \frac{1 - \varrho_{ij}}{\sum_{x=1}^{n_i^k} (1 - \varrho_{ix})} \quad (3)$$

The first term of the equation captures the contribution of the highest priority VPs that are busy for commodity k . The second term considers the case where all the forward VPS are busy and the link (i, j) becomes the first *empty* one.

2.2 Distance Measures for Commodity Switching

A first measure of distance which is called the *expected proximity* (EP) of the node i to node j for the commodity k , and is defined as follows

$$EP_i^{k,j} = \sum_{p \in P} \frac{1}{|p|} \prod_{(u,v) \in p} (1 - \varrho_{uv}) \quad (4)$$

where P denotes all possible forward paths from i to j for commodity k on the routing table, $|p|$ is the length of a path measured in hops, and (u, v) is a link between nodes u and v . The measure $EP_i^{k,j}$ has relatively large values when there are few paths p between i and j , but they are short and lightly loaded or there exists a large number of paths between i and j .

A second measure is the *maximum proximity* (MP) of node i to node j for the commodity k is which is defined as follows

$$MP_i^{k,j} = \max_{p \in P} \left\{ \frac{1}{|p|} \prod_{(u,v) \in p} (1 - \varrho_{uv}) \right\} \quad (5)$$

The maximum proximity provides a finer distinction of the nodes that have better forward paths to destination.

3 Markov Chain Approach to Traffic Class Migration

3.1 Markov Processes for Dynamic Modeling

Definition 3.1. (Markov Modulated Poisson Process (MMPP)). Let X_n be a n -state Markov chain with matrix of transition probabilities $\mathbf{Q}$. Doubly stochastic Poisson process is called Markov modulated Poisson process (MMPP) with vector of Poisson arrival rates $\{\lambda_i\}_{1 \leq i \leq n}$ if its arrival rate is λ_i when underlying Markov chain X_n is in the state i . $\square$

Remark 3.1. MMPP with n -state is described by a pair of parameters $\Lambda = \{\lambda_i\}_{1 \leq i \leq n}$ and matrix of transition probabilities $\mathbf{Q}$. The probability q_{ij} denotes the probability that the chain, whenever in the state i , moves next (one unit of time later) into state j and can be referred to as a one-step transition probability. The square matrix $\mathbf{Q} = [q_{ij}]_{i \in S, j \in S}$, $S = \{1, \dots, n\}$ is called the one-step transition matrix. Obviously, the following normalization constraint should be satisfied

$$\sum_{j \in S} q_{ij} = 1 \quad (6)$$

$\square$

In order to describe the user class dynamic behaviour via MMPP models for discrete Markov chain, let us review the MMPP driven by a continuous time Markov chain. The differential equation formalism is an important step towards local scale characterization of events of the stochastic systems represented by the communication networks.

Definition 3.2. (Continuous Time Markov Chain (CTMC)). A stochastic process $\mathbf{X} = \{X_t\}$, $t > 0$ is continuous Markov chain with states in S if

$$q_{ij}(s) = \Pr[X(t+s) = j | X(t) = i], \forall i, j \in S, \forall s > 0 \quad (7)$$

$\square$

Proposition 3.1. The Chapman-Kolmogorov equation states that

$$\mathbf{Q}(s+t) = \mathbf{Q}(s)\mathbf{Q}(t), \forall s > 0, t > 0 \quad (8)$$

where $\mathbf{Q}(0) = \mathbf{I}$. If we set $t = ds$, we obtain the following equation

$$\frac{d}{ds}\mathbf{Q}(s) = \mathbf{Q}(s)\mathbf{G} \quad (9)$$

where $\mathbf{G} = \frac{d}{ds}\mathbf{Q}(s)|_{s=0}$ is the generator matrix of the Markov chain. $\square$

In order to relate the continuous time Markov chain models to the traffic dynamic behaviour, let $N(t)$ be the number of arrivals in an interval $(0, t]$ and J_t be the state of the Markov chain at the time step t . Define the matrix $\mathbf{Q}(n, t)$, $n \in S$, $t > 0$ with elements $[q_{ij}(n, t)]_{i \in S, j \in S}$ as follows

$$q_{ij}(n, t) = \Pr[N_t = n, J_t = j | N_0 = 0, J_0 = i] \quad (10)$$

We call the matrix $\mathbf{Q}(\mathbf{n}, t)$ the probability transition matrix of double process (N_t, J_t) . Let Λ be the diagonal matrix associated with vector of intensities of MMPP, that is, $\Lambda = \text{diag}[\lambda_i]_{i \in S}$. The matrices $\mathbf{Q}(n, t)$, $n \in S$, $t > 0$, satisfy the *Chapman-Kolmogorov* equations

$$\frac{d}{dt} \mathbf{Q}(n, t) = \mathbf{Q}(n, t)(\mathbf{G} - \Lambda) + \mathbf{Q}(n-1, t)\Lambda \quad (11)$$

$$\mathbf{Q}(0, 0) = \mathbf{I}, \quad \forall n \in S, \quad t > 0 \quad (12)$$

Multiplying (11) by z^n and summing for $n = 0, 1, \dots$, we obtain

$$\frac{d}{dt} Z\{\mathbf{Q}(z, t)\} = Z\{\mathbf{Q}(z, t)\}(\mathbf{G} - \Lambda) + zZ\{\mathbf{Q}(z, t)\}\Lambda \quad (13)$$

$$Z\{\mathbf{Q}(z, t)\}|_{t=0} = \mathbf{I} \quad (14)$$

where $Z\{\mathbf{Q}(z, t)\} = \sum_{n=0}^{\infty} z^n \mathbf{Q}(n, t)$ is the Z -transform of matrix $\mathbf{Q}(n, t)$. Solving (13), we get the generating function of $\mathbf{Q}(n, t)$ as follows

$$Z\{\mathbf{Q}(n, t)\} = \exp[\mathbf{G} + (z-1)\Lambda t] \quad (15)$$

The expected number of arrivals number in the interval $(0, t]$ can be derived from the expression of the generating function, that is,

$$\mathbf{E}[N_t] = \pi \frac{\partial Z\{\mathbf{Q}(n, t)\}}{\partial z} \Big|_{z=1} \mathbf{e} \quad (16)$$

where $\mathbf{e} = [1, \dots, 1]^T$ and π is the steady state vector of the Markov chain such that $\pi \mathbf{Q} = \pi$. Finally, we obtain

$$\mathbf{E}[N_t] = \pi \Lambda \mathbf{e} t = \sum_{s \in S} \pi_s \lambda_s t \quad (17)$$

Remark 3.2. The component π_i of steady state vector π can be interpreted as asymptotically reached probability to be in the state $i \in S$ without regard of the history, or it can be expressed as

$$\lim_{t \rightarrow \infty} \Pr[X(t) = j | X(0) = i] = \pi_i, \quad \forall i, j \in S \quad (18)$$

□

The type of developments presented in this section is widely used to characterize the queueing processes associated to the routing and switching device in the communication network.

3.2 Markov Chain Description of Traffic Migration

Let us consider that the traffic sources dynamics which are characterized by a MMPP process are subject to changes their operation regimes. We will call this feature, traffic class migration, and we attempt to use the available results from the theory of discrete Markov chains in order to encode the most significant feature of their behaviour.

Let us denote τ_j the first passage time to state j , defined as $\tau_j = \min \{n \geq 1, X(n) = j\}$. The probability distribution of τ_j is of great importance in what follows. Let us define

$$f(n, i, j) = \Pr_i(\tau_j = n), \quad n \geq 1 \quad (19)$$

and

$$f(i, j) = \sum_{n \geq 1} f(n, i, j) = \Pr_i(\tau_j < \infty) \quad (20)$$

where

$$\Pr_i[\tau = \infty] = f(\infty, i, j) = 1 - f(i, j) \quad (21)$$

Therefore, $f(n, i, j)$, $n = 1, 2, \dots, \infty$, is the probability distribution of τ_j under the probability $\Pr_i$ (i.e., given the Markov chain starts in state i). The probabilities considered can be obviously expressed in a way that involves the Markov chain itself. Then, we have

$$\{\tau_j = 1\} = \{X(1) = j\} \quad (22)$$

$$\{\tau_j = n\} = \{X(m) \neq j, 1 \leq m \leq n-1, X(n) = j\}, \quad n \geq 2 \quad (23)$$

$$\{\tau_j = \infty\} = \{X(m) \neq j, m \geq 1\} \quad (24)$$

$$\{\tau_j < \infty\} = \{X(m) = j \text{ for at least one value of } m \geq 1\} = \bigcup_{m \geq 1} \{X(m) = j\}, \quad (25)$$

Therefore

$$f(1, i, j) = \Pr_j[X(1) = j] = p(i, j) \quad (26)$$

$$f(n, i, j) = \Pr_j[X(m) \neq j, 1 \leq m \leq n-1, X(n) = j] \quad (27)$$

$$= \sum_{i_m \neq j, (1 \leq m \leq n-1)} p(i, i_1) p(i_1, i_2) \dots p(i_{n-1}, j), \quad n \geq 2 \quad (28)$$

$$f(\infty, i, j) = \Pr_i[X(m) \neq j, m \geq 1] \quad (29)$$

$$f(i, j) = \Pr_i[X(m) = j \text{ for at least one value of } m \geq 1] = \Pr_i\left[\bigcup_{m \geq 1} \{X(m) = j\}\right] \quad (30)$$

The last equalities lead to

$$\Pr_i[X(n) = j] \geq f(i, j) \geq \sum_{m \geq 1} \Pr_i[X(m) = j] \quad (31)$$

for all $n \geq 1$, where

$$\sup_{n \geq 1} p(n, i, j) \geq f \geq \sum_{m \geq 1} p(m, i, j) \quad (32)$$

for all states i and j . It follows at once that $i \rightarrow j$ if and only if $f(i, j) > 0$ and that $i \leftrightarrow j$ if and only if $f(i, j)f(j, i) > 0$. Notice that by homogeneity we also have

$$f(1, i, j) = \Pr[X(m+1) = j | X(m) = i] \quad (33)$$

$$f(n, i, j) = \Pr[X(m+l) \neq j, 1 \leq l \leq n-1, X(m+n) = j | X(m) = i], n \geq 2 \quad (34)$$

for any m for which the conditional probabilities are defined. The following results describe the qualitative properties of the Markov chain that are relevant in describing the traffic migration processes.

Theorem 3.1. *If the Markov chain starts in state i , then the probability of returning to i at least r times equals $[f(i, i)]^r$.* $\square$

Corollary 3.1. *Assume the Markov chain starts in state i . If $f(i, i) = 1$, then probability of returning to i infinitely often is 1. If $f(i, i) < 1$, the probability of returning to i infinitely often is 0.*

Proof. The random event

$$A = \{ \text{the Markov chain returns to } i \text{ infinitely often} \} \quad (35)$$

is the intersection of the decreasing sequence of random events

$$\{ \text{the Markov chain returns to } i \text{ at least } r \text{ times} \}, r \geq 1 \quad (36)$$

Therefore, the probability of A is equal to

$$\lim_{m \rightarrow \infty} [f(i, i)] = \begin{cases} 1, & \text{if } f(i, i) = 1, \\ 0, & \text{if } f(i, i) < 1. \end{cases} \quad (37)$$

$\square$

Theorem 3.2. (First Entrance Theorem). *Whatever the states i and j and the natural number n , we have*

$$p(n, i, j) = \sum_{m=1}^n f(m, i, j) p(n-m, j, j) \quad (38)$$

Proof. Intuitively we may argue as follows. To be in state j at the n th step, the Markov chain should reach that state for the first time at some time m , $1 \leq m \leq n$. After that happens, it should return to j in $n - m$ steps. The rigorous proof uses the strong Markov property. We have

$$p(n, i, j) = \Pr_i[X(n) = j] = \Pr_i[\tau_j \leq n, X(n) = j] \quad (39)$$

since the random event $\{X(n) = j\}$ implies the random event $\{\tau_j \leq n\}$

$$p(n, i, j) = \sum_{m=1}^n \Pr_i[\tau_j = m, X(n) = j] \quad (40)$$

Because $\{\tau_j \leq n\}$ is the union of disjoint random events $\{\tau_j = m\}$, $1 \leq m \leq n$

$$p(n, i, j) = \sum_{m=1}^n \Pr_i[\tau_j = m, X(\tau_j + n - m) = j] \quad (41)$$

Since $\{\tau_j = m\} = \{\tau_j + n - m = n\}$

$$p(n, i, j) = \sum_{m=1}^n \Pr_i[\tau_j = m] \Pr_i[X(\tau_j + n - m) = j | \tau_j = m] \quad (42)$$

$$p(n, i, j) = \sum_{m=1}^n \Pr_i[\tau_j = m] \Pr_i[X(\tau_j + n - m) = j | X(\tau_j) = j, \tau_j = m] \quad (43)$$

By Corollary 3.1 of Theorem 3.1, then we have

$$p(n, i, j) = \sum_{m=1}^n f(m, i, j) p(n - m, j, j) \quad (44)$$

□

Theorem 3.3. (Doebelin's formula). For any states i and j , we have

$$f(i, j) = \lim_{s \rightarrow \infty} \frac{\sum_{n=1}^s p(n, i, j)}{1 + \sum_{n=1}^s p(n, j, j)} \quad (45)$$

Proof. Equation (38) yields

$$\sum_{n=1}^s p(n, i, j) = \sum_{n=1}^s \sum_{m=1}^n f(m, i, j) p(n - m, j, j) = \sum_{m=1}^s (f(m, i, j) \sum_{n=m}^s p(n - m, j, j)), \quad (46)$$

Thus,

$$\left(1 + \sum_{n=1}^s p(n, j, j)\right) \sum_{m=1}^s f(m, i, j) \geq \sum_{n=1}^s p(n, i, j) \geq \left(1 + \sum_{n=1}^{s-s'} p(n, j, j)\right) \sum_{m=1}^{s'} f(m, i, j), \quad (47)$$

for all $s' < s$. Since $1 + \sum_{n=1}^s p(n, j, j)$ dominates

$$\sum_{n=m}^s p(n-m, j, j) = 1 + \sum_{n=1}^{s-m} p(n, j, j) \quad (48)$$

and

$$\sum_{m=1}^s (f(m, i, j) \sum_{n=m}^s p(n-m, j, j)) \quad (49)$$

dominates

$$\sum_{m=1}^{s'} \left((f(m, i, j) \sum_{n=m}^{s-s'+m} p(n-m, j, j)) \right) = \left(1 + \sum_{n=1}^{s-s'} p(n, j, j) \right) \sum_{m=1}^{s'} f(m, i, j). \quad (50)$$

dividing by $1 + \sum_{n=1}^s p(n, j, j)$ and letting first $s \rightarrow \infty$, then $s' \rightarrow \infty$, yields the equation (45). $\square$

Remark 3.3. The above results are useful towards estimating theoretically the transition rates of the discrete Markov chain which is modulating the basic Poisson process. An extension of these results in the spirit of properties presented in Section 3.1 would be of major interest. $\square$

4 Input-Output Representation of Statistical Behaviour

Let us consider an extended experiment that can be formulated as follows. Assume that for each commodity $w \in W$, we have the set of possible values for vector $\Lambda_w \in \mathcal{R}^n$ denoted by $\mathcal{L}_{in} \subset \mathcal{R}^n$. The estimated set of parameters is $\hat{\mathcal{L}}_{in} \subset \mathcal{R}^n$. Similarly, $\hat{\mathcal{L}}_{dep} \subset \mathcal{R}^n$ is the set of intensity parameters for the output traffic. The goal is to find the dependency relation between the sets $\hat{\mathcal{L}}_{in}$ and $\hat{\mathcal{L}}_{dep}$. Let $\hat{\Lambda}_{in} \in \hat{\mathcal{L}}_{in}$ and $\hat{\Lambda}_{dep} \in \hat{\mathcal{L}}_{dep}$.

$$\hat{\Lambda}_{dep} = f(\hat{\Lambda}_{in}) \quad (51)$$

We concentrate on constructing linear relations as follows

$$\hat{\Lambda}_{dep} = \mathbf{G}_{LS} \hat{\Lambda}_{in} \quad (52)$$

where the equality holds in the least squares sense [5], that is,

$$\mathbf{G}_{LS} = \arg \min_{\mathbf{G} \in \mathcal{R}^{n \times n}} \sum_{\hat{\Lambda}_{in} \in \hat{\mathcal{L}}_{in}, \hat{\Lambda}_{dep} \in \hat{\mathcal{L}}_{dep}} \|\hat{\Lambda}_{dep} - \mathbf{G} \hat{\Lambda}_{in}\|^2 \quad (53)$$

Finding the matrix $\mathbf{G}_{LS}$ will encode the dependencies between the statistical representations of the input and output traffic. Instead of proposing the model for the switching system and then estimating the model parameters, we consider models for input and the output traffic and in the extended experiment, the changes in the output parameters will reflect changes in the input parameters that are subject of statistical estimation [2].

The matrix $\mathbf{G}_{LS}$ can be used further in some other ways. First, the explicit usage is in the prediction of the output parameters for given vectors of input parameters which in fact may not belong to the set $\hat{\mathcal{L}}_{in}$ of extended experiment. In this case, it can be seen as an extrapolation, or as interpolation if a given vector is included into $\hat{\mathcal{L}}_{in}$. In both cases, this extended experiment may be treated as a learning process during which the matrix $\mathbf{G}_{LS}$ is being constructed. In the next stage, $\mathbf{G}_{LS}$ is used without modifications as a given model for dependencies between the input and the output. The matrix $\mathbf{G}_{LS}$ can be estimated as off-line linear regression over the sets $\hat{\mathcal{L}}_{in}$ and $\hat{\mathcal{L}}_{dep}$.

The second possible usage of $\mathbf{G}_{LS}$ is in solving the inverse problem of finding the values of the input parameters corresponding to given output parameters. This problem can be solved in the context of adaptive control (or regulation) of the switching system in order to achieve a desired level of performance with respect to given regulation criteria, which in the most situations can be expressed as a desired set of output traffic parameters.

5 Derivation of GLR Test

In this section we derive GLR algorithm that is effectively used under the Poisson distribution hypothesis for the basic stochastic process of traffic generation [1]. We consider a parametric family of Poisson distributions $\mathcal{P} = \{\text{Pr}_{\lambda}\}_{\lambda \in \Theta}$, $\Theta \subset \mathbb{R}^n$. Let as usual, $\mathcal{Y}_1^N$ be a data sample of size N . The goal is to distinguish the distance between two hypothesis

$$\begin{aligned} \mathbf{H}_0 &= \{L = \text{Pr}_{\lambda_0}\} \\ \mathbf{H}_1 &= \{L = \text{Pr}_{\lambda_1}\} \end{aligned} \quad (54)$$

Since the Poisson process can be represented as the sum of i.i.d. exponentially distributed random variables, its density is given by

$$p_{n,\lambda}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, \quad t \geq 0 \quad (55)$$

The log-likelihood ratio of the data sample $\mathcal{Y}_1^N$ for observations from time j till N plays the crucial role in the construction of GLR method and it is defined as follows

$$S_j^N(\lambda_0, \lambda_1) = \ln \frac{p_{N,j,\lambda_1}(\mathcal{Y}_j^N)}{p_{N,j,\lambda_0}(\mathcal{Y}_j^N)} \quad (56)$$

where $p_{N,j,\lambda_i}(\mathcal{Y}_j^N)$, $i = 1, 2$ is the joint p.d.f. for part of the data sample for time steps $j, j+1, \dots, N$.

Remark 5.1. In the notation of joint p.d.f., there are subscripts N and j that explicitly denote the dependency p.d.f. of time interval of considered part of N -size sample $\mathcal{Y}_1^N$. $\square$

Under the independence assumption of observations, equation (56) becomes

$$S_j^N(\lambda_0, \lambda_1) = \sum_{t=j}^N \ln \frac{p_{n,\lambda_1}(y_t)}{p_{n,\lambda_0}(y_t)} \quad (57)$$

Let us derive the expression for joint p.d.f of $\mathcal{Y}_1^N$. We have

$$\begin{aligned} p_{N,j,\lambda}(\mathcal{Y}_j^N) &= \prod_{i=j}^N \lambda e^{-\lambda y_i} \frac{(\lambda y_i)^{(i-1)}}{(i-1)!} \\ &= \lambda^{N-j+1} e^{-\lambda(\sum_{i=j}^N y_i)} \frac{\lambda^{(j-1)+j+\dots+N}}{(j-1)!j!\dots N!} \prod_{i=j}^N y_i^{(i-1)} \\ &= \pi \lambda^\alpha e^{-\lambda\beta} \end{aligned} \quad (58)$$

where in the last relation, the following notations are used

$$\pi = \prod_{i=j}^N y_i^{(i-1)} \quad (59)$$

$$\alpha = N - j + 1 + \frac{1}{2}(N(N-1) - (j-1)(j-2)) \quad (60)$$

$$\beta = \sum_{i=j}^N y_i \quad (61)$$

The maximum of $p_{N,j,\lambda}(\mathcal{Y}_j^N)$ with respect to the parameter λ can be easily obtained after solving the equation

$$\frac{\partial p_{N,j,\lambda}(\mathcal{Y}_j^N)}{\partial \lambda} = 0 \quad (62)$$

Let us explicitly solve the equation (62) which can be rewritten as

$$\frac{\partial p_{N,j,\lambda}(\mathcal{Y}_j^N)}{\partial \lambda} = \pi [\alpha \lambda^{(\alpha-1)} e^{-\lambda\beta} - \lambda^\alpha \beta e^{-\lambda\beta}] = 0 \quad (63)$$

Consequently,

$$\hat{\lambda}(N, j, \mathcal{Y}_1^N) = \arg \sup_{\lambda} [p_{N,j,\lambda}(\mathcal{Y}_j^N)] = \frac{\alpha - 1}{\beta} \quad (64)$$

In order to formulate GLR method in the case of *unknown* parameter λ_1 after change, the equation (56) of the log-likelihood ratio should be rewritten as

$$S_j^N(\lambda_0) = \ln \sup_{\lambda_1} \left[\frac{p_{N,j,\lambda_1}(\mathcal{Y}_j^N)}{p_{N,j,\lambda_0}(\mathcal{Y}_j^N)} \right] \quad (65)$$

which using (64) can be written as follows

$$S_j^N(\lambda_0) = \ln \left[\frac{p_{N,j,\hat{\lambda}(N,j,\mathcal{Y}_1^N)}(\mathcal{Y}_j^N)}{p_{N,j,\lambda_0}(\mathcal{Y}_j^N)} \right] \quad (66)$$

$$\begin{aligned} S_j^N(\lambda_0) &= \ln \left[\left(\prod_{i=j}^N \hat{\lambda} e^{\hat{\lambda} y_i} \frac{(\hat{\lambda} y_i)^{(i-1)}}{(i-1)!} \right) \left(\prod_{i=j}^N \lambda_0 e^{\lambda_0 y_i} \frac{(\lambda_0 y_i)^{(i-1)}}{(i-1)!} \right)^{-1} \right] \\ &= \ln \left[\left(\frac{\hat{\lambda}}{\lambda_0} \right)^{N-j} \frac{e^{(N-j)\hat{\lambda}} \hat{\lambda}^{(j-1)+j+\dots+N}}{e^{(N-j)\lambda_0} \lambda_0^{(j-1)+j+\dots+N}} \right] \\ &= (N-j + \sum_{i=j-1}^N) \ln \left(\frac{\hat{\lambda}}{\lambda_0} \right) + (N-j)(\hat{\lambda} - \lambda_0) \end{aligned} \quad (67)$$

where $\hat{\lambda} = \hat{\lambda}(N, j, \mathcal{Y}_1^N)$ and is given by (64). Finally, the GLR statistics are given by

$$g_k(S_1^k) = \max_{j \in [1, k]} S_j^k(\lambda_0), \quad k > 0 \quad (68)$$

The stopping time of GLR test can be written as

$$t_a^{GLR} = \min_{k > 0} \{g_k(S_1^k) \geq h_{GLR}\} \quad (69)$$

The threshold h_{GLR} is an adjustable parameter of the proposed test. The choice of this threshold is based on the investigation of asymptotic behaviour of the likelihood ratio. Let α be a size of test $g_k(S_1^k)$ when the asymptotic distribution of $-2 \ln g_k(S_1^k)$ under hypothesis $\mathbf{H}_0$ is $\chi^2(1)$. The following relation determines the *in-control state* of $\mathcal{Y}_1^N$

$$-2 \ln g_k(S_1^k) \leq \chi_{1-\alpha}^2(1) \quad (70)$$

where $\chi_{1-\alpha}^2(1)$ is the $(1 - \alpha)$ -th quantile of the χ^2 -distribution with 1 degree of freedom.

Remark 5.2. The same asymptotic behaviour of $g_k(S_1^k)$ can be observed in the vector case of parameter. In this case

$$L[-2 \ln g_k(S_1^k)] \Rightarrow \chi^2(l), \quad \text{when } L(S_1^k) = \mathbf{P}_{\lambda_0} \quad (71)$$

where l is the dimension of parameter space. $\square$

6 Concluding Remarks

In this paper we have shown the main concepts from the telecommunication networks traffic theory and system modelling and their interplay that are useful in order to describe in an unified manner the input-output transfer processes occurring in the switching systems. Some experimental studies are to be considered in the future towards validating the theoretical framework that has been developed.

References

- [1] M. Basseville, I.V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, Englewoods Cliffs, NJ, 1993.
- [2] Y. Bengio, P. Frasconi, "Input/Output HMMs for Sequence Processing", *Research Report*, AT&T Bell Laboratories, Holmdel, NJ, September 4, 1995.
- [3] M. Iosifescu, *Finite Markov Processes and Their Applications*, John Wiley and Sons, New York, 1980.
- [4] D. Medhi, "Multi-Hour, Multi-Traffic Class Network Design for Virtual Path-Based Dynamically Reconfigurable Wide-Area ATM Networks", *IEEE/ACM Transactions on Networking*, vol.3, no. 6, pp. 809-818, 1995.
- [5] E. Mosca, *Optimal, Predictive and Adaptive Control*, Prentice-Hall, Englewoods-Cliffs, NJ, 1995.
- [6] A. Murgu, *Optimization of Telecommunication Networks. Lecture Notes*, University of Jyväskylä, 1999.
- [7] J. Rittwik, D. Subhrakanti, "Change Detection in Teletraffic Models", in *Proceedings of the 36th IEEE CDC*, San Diego, pp. 3984-3989, 1997.

СОДЕРЖАНИЕ

<i>А.Ю. Богданов, В.В. Богданов</i> Моделирование вероятности образования среза с учетом радиусов закругления эллипсоидальных зерен шлифовального круга и их ориентации .	3
<i>А.Г. Бросалина</i> Об условиях коммутирования в глобальном надмоноиде свободного моноида .	6
<i>М.В. Бузин</i> Методы единого поиска на примере игры «Сокобан» .	22
<i>И.Р. Васильева</i> Критерий конечности кодлины многообразий Z_2 -градуированных ассоциативных алгебр .	35
<i>П.А. Вельмисов, С.В. Киреев, А.О. Кузнецов.</i> Асимптотика решений задачи об устойчивости пластины в сверхзвуковом потоке газа .	41
<i>Г.М. Ильмушкин</i> Об одной формуле обобщенных резольвент изометрических операторов .	49
<i>П.А. Новиков</i> Задание кусочно-непрерывных функций и действия над ними .	58
<i>Н.В. Романов</i> Система мультиагентной оптимизации для задачи коммивояжера .	63
<i>Ю.А. Тимаков</i> Полилинейные тождества матричных супералгебр типа $M_k \otimes G_n$ над полем характеристики нуль .	75
<i>Л.А. Штраус</i> Спектральное представление вполне неунитарного оператора .	84
<i>Ю.В. Цыганова, А.Д. Юрьев</i> Эффективные инструментальные средства моделирования линейных стохастических систем управления .	89
<i>И.В. Семушин, Ю.В. Цыганова, М.В. Куликова</i> О вычислении функции правдоподобия для гауссовских марковских последовательностей .	93
<i>A. Murgu, O. Yu. Gorokhov, V. Semoushin</i> Input-Output Statistical Inference for Switching Processes .	101

УЧЕННЫЕ ЗАПИСКИ
Ульяновского государственного университета

Серия
Фундаментальные проблемы
математики и механики

Вып.2(9)

Под ред. академ. РАН, проф. *А.С.Андреева*

Оригинал-макет подготовил *Е.В.Дулов*

Подписано в печать 22.12.2000. Формат 60x84/8.
Бумага книжно-журнальная. Усл. печ.л. 13,11. Уч.-изд. л. 7,32.
Тираж 100 экз. Заказ №167/**317**

Отпечатано с оригинал-макета в лаборатории оперативной полиграфии
Ульяновского государственного университета
432700, г.Ульяновск, ул.Л.Толстого, 42

THE UNIVERSITY OF CHICAGO
PRESS

