

СРАВНЕНИЕ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ АНАЛИЗА СООТВЕТСТВУЮЩИХ ГРАФОВ

С.Н. Чуканов¹

д.т.н., профессор, ведущий научный сотрудник, e-mail: ch_sn@mail.ru

И.С. Чуканов²

магистрант, e-mail: chukanov022@gmail.com

С.В. Лейхтер³

старший преподаватель, e-mail: leykhter@mail.ru

¹Институт математики им. С.Л. Соболева СО РАН, Омский филиал, Омск, Россия

²Уральский федеральный университет имени первого Президента России Б.Н. Ельцина,
Екатеринбург, Россия

³Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Рассматривается сравнение временных рядов совокупности концепций на основе формирования взвешенных ориентированных графов взаимных влияний концепций временных рядов друг на друга; при этом вершинам графа соответствуют концепции, а взвешенным дугам значения величин влияния одной концепции на другую. Структуру графа можно охарактеризовать с помощью спектра собственных значений матрицы Лапласа графа, связанного с уравнением теплопроводности. Для этого формируется матрица теплового ядра; производная матрицы теплового ядра по времени определяется матрицей Лапласа. Коэффициенты разложения теплового содержания теплового ядра в ряд Тейлора можно использовать для представления структуры графа. В работе представлены соотношения для определения расстояния Махаланобиса между векторами признаков – инвариантами теплового ядра, которые могут быть использованы для распознавания образов. Актуальность проблемы – повышение уникальности идентификации при распознавании временных рядов за счет учета как спектра собственных значений матрицы Лапласа графа, так и собственных векторов. Цель исследования – выяснить возможность использования коэффициентов степенного ряда разложения теплового содержания построенного графа в качестве векторов признаков – характеристик свойств графа. Новизна представленных в работе методов заключается в том, что, в отличие от метода спектрального вложения, в работе исследуется возможность сравнения графов (изображений) на основе расстояния Махаланобиса между инвариантами, построенными по коэффициентам разложения теплового содержания. Предлагается алгоритм минимизации расстояния Махаланобиса между вершинами графа, основанный на обучении матрицы Махаланобиса. На основе сравнения графов, соответствующих временным рядам совокупности концепций, можно сравнивать временные ряды.

Ключевые слова: тепловое ядро, расстояние Махаланобиса, временные ряды.

Введение

В работе рассматривается сравнение временных рядов совокупности концепций на основе формирования взвешенных ориентированных графов взаимных влияний концепций временных рядов друг на друга; при этом вершинам графа соответствуют концепции, а взвешенным дугам значения величин влияния одной концепции на другую. Структуру графа можно охарактеризовать с помощью спектра собственных значений матрицы Лапласа графа [1]. Спектр матрицы Лапласа связан с уравнением теплопроводности [2], которое может быть использовано для определения потока информации через сеть графа. Производная теплового ядра по времени определяется матрицей Лапласа. Тепловое содержание (сумма множества элементов теплового ядра) может быть разложено в степенной ряд по времени [4]. Так как решение уравнения теплопроводности является функцией времени, то в [3] рассматривается алгоритм определения матрицы теплового ядра для конкретного момента времени; выбор конкретного момента времени не указывается. В настоящей работе рассматривается определение коэффициентов разложения теплового содержания (сумма элементов матрицы теплового ядра по узлам графа) при разложении теплового содержания в степенной ряд по времени. Коэффициенты разложения теплового содержания теплового ядра в ряд Тейлора можно использовать для представления структуры графа. Для сравнения графов предлагается находить расстояние Махаланобиса [5, 6] между векторами признаков – коэффициентами разложения теплового содержания. Определение расстояния Махаланобиса между векторами признаков и элементарными симметрическими многочленами лапласианов графов может быть использовано для распознавания формы объектов. Цель работы – выяснить возможность использования коэффициентов степенного ряда разложения теплового содержания построенного графа в качестве векторов признаков – характеристик свойств графа. Новизна представленных в работе методов заключается в том, что, в отличие от метода спектрального вложения, в работе исследуется возможность сравнения графов (изображений) на основе расстояния Махаланобиса между инвариантами, построенными по коэффициентам разложения теплового содержания. Предлагается алгоритм минимизации расстояния Махаланобиса между вершинами графа, основанный на обучении матрицы Махаланобиса. На основе сравнения графов, соответствующих временным рядам совокупности концепций, можно сравнивать временные ряды. Применение метода сравнения временных рядов основе построения и сравнения графов позволит идентифицировать штатное и нештатное функционирование динамических объектов, систем управления, приборного состава и т.д.

1. Тепловое ядро на графах

Пусть $G = \{V, E, W\}$ – неориентированный взвешенный граф, где $V = \{v_1, \dots, v_n\}$ – множество узлов графа, представленных в евклидовых координатах. $E = \{e_1, \dots, e_m\}$ – множество неориентированных рёбер, $W = [w_{ij}]$; $i, j \in \{1, \dots, n\}$ – квадратная взвешенная матрица смежности графа. Матрица Ла-

матрицы L графа может быть определена как:

$$L = D - W, \quad (1)$$

где $D \in \mathbb{R}^{n \times n}$ – диагональная матрица с элементами $D_{ii} = \sum_{j=1}^n w_{ij}$. Вложение узлов графа в евклидово пространство может быть получено с помощью собственного спектра матрицы Лапласа графа. Пусть $L = U \Lambda U^T$ – это спектральное разложение матрицы L , где $U = [u_1, \dots, u_n]$ – собственные векторы, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ – собственные значения матрицы Лапласа L . Уравнение теплопроводности, связанное с матрицей Лапласа, может быть записано как:

$$\partial h_t / \partial t = -L h_t, \quad (2)$$

где h_t – тепловое ядро, соответствующее графу. Решение уравнения теплопроводности имеет вид: $h_t = \exp(-tL) \in \mathbb{R}^{n \times n}$. Определим тепловое ядро на графе на основе спектра Лапласа [3]:

$$h_t(i, j) = \Phi \exp(-\Lambda t) \Phi^T = \sum_{l=1}^n \exp(-\lambda_l t) u_{li} u_{lj}^T. \quad (3)$$

Тепловое содержание – это сумма элементов матрицы теплового ядра по узлам графа:

$$Q(t) = \sum_i \sum_j h_t(i, j) = \sum_i \sum_j \sum_{k=1}^n e^{-\lambda_k t} u_{ik} u_{jk}. \quad (4)$$

Тепловое содержание можно разложить в степенной ряд по времени [4]:

$$Q(t) = \sum_{m=0}^{\infty} q_m t^m. \quad (5)$$

После разложения экспоненты $e^{-\lambda_k t}$ в ряд Тейлора коэффициенты q_m разложения $Q(t)$ можно представить следующим образом:

$$q_m = \sum_{i,k} \frac{(-\lambda_k)^m}{m!} \left(\sum_i \sum_j u_{ik} u_{jk} \right) = \sum_{k=1}^n \frac{(-\lambda_k)^m}{m!} \left(\sum_i u_{ik} \right)^2. \quad (6)$$

Коэффициенты q_m зависят от собственных значений и собственных векторов; они более уникальны, чем инварианты, основанные только на собственных значениях. Пусть q_1 – вектор коэффициентов разложения $Q(t)$, построенного по лапласиану для графа G^1 ; q_2 – вектор коэффициентов, построенного по лапласиану для графа G^2 . Расстояние Махаланобиса (взвешенное евклидово расстояние) между векторами коэффициентов q_1 и q_2 , позволяющее различать графы G^1 и G^2 , можно определить по формуле [5, 6]:

$$d(q_1, q_2)^2 = \text{trace} \left((q_1 - q_2) W_q (q_1 - q_2)^T \right), \quad (7)$$

где $W_q = \text{diag} \begin{pmatrix} a_1 & a_2 & a_3 & \dots \end{pmatrix}$ – диагональная матрица Махаланобиса положительных весов. Поставим задачу обучения матрицы W_q , основанную на минимизации расстояния $d(q_1, q_2)^2$ путём выбора матрицы

$$W_q : \min_A \text{trace} \left((q_1 - q_2) W_q (q_1 - q_2)^T \right), \quad (8)$$

при ограничении: $\sum_i (W_q)_i = 1$.

2. Сравнение временных рядов

Нечеткие когнитивные карты состоят из концепций $C_i, i = 1, \dots, n$. Каждый узел (концепция) содержит значение $A_i \in [0, 1], i = 1, \dots, n$. Концепции соединены взвешенными ориентированными дугами, которые показывают отношения между концептами. Дуга между концепциями A_i и A_j имеет вес w_{ij} . Направление указывает, является ли концепция причиной или следствием. На каждом шаге значение концепции A_i находится под влиянием значений других концепций:

$$A_i(k+1) = f \left(A_i(k) + \sum_{\substack{j=1 \\ j \neq i}}^n w_{ji} A_j(k) \right), i = 1, \dots, n, \quad (9)$$

где w_{ij} – вес дуги, соединяющей концепцию причины C_j с концепцией следствия C_i ; $f(x)$ – сигмоидальная функция: $f(x) = (1 + \exp(-\lambda(x - \varsigma)))^{-1}$, где λ – значение крутизны сигмоидальной функции, ς – значение, при котором $f(\sigma) = 0.5$. Процесс обучения используется для определения матрицы весов $W = [w_{ij}]$. Доступные для обучения данные поступают в виде последовательности векторов $A_c(k), k = 1, 2, \dots, N$. Функция критерия, по экстремуму (минимуму) которой определяются требуемые значения элементов матрицы W , определяется как сумма расстояний между прогнозируемыми векторами активации узлов $A(k+1)$, которые вычисляются на основе $A(k)$, доступных в k -й момент времени, и векторами $\tilde{A}(k+1)$ известных уровней активации узлов $(k+1)$ -й момент времени:

$$V = \min_{w_{ij}} \left(\sum_{k=1}^N \sum_{i=1}^n \left\| \tilde{A}_i(k+1) - f_i \left(\sum_{j=1}^n w_{ij} A_j(k) \right) \right\|^2 \right). \quad (10)$$

Рассмотрим пример оценивания весов взаимного влияния концепций экологии города [7, 8] по информации о временных рядах этих концепций. Введем следующие обозначения концепций: A_1 – население города; A_2 – миграция в город; A_3 – модернизация; A_4 – мусор на единицу площади; A_5 – санитарные условия; A_6 – количество заболеваний; A_7 – количество бактерий на единицу площади. В таблице 1 заданы исходные временные ряды концепций $A_1 - A_7$ при компонентах матрицы взаимных

Таблица 1. Исходный временной ряд

t	A1	A2	A3	A4	A5	A6	A7
0	0.300	0.350	0.400	0.450	0.500	0.550	0.600
1	0.166	0.634	0.579	0.760	0.872	0.597	0.568
2	0.098	0.934	0.715	0.890	0.988	0.236	0.460
3	0.132	0.991	0.774	0.903	0.997	0.019	0.345
4	0.208	0.994	0.837	0.941	0.997	0.004	0.232
5	0.283	0.996	0.907	0.963	0.998	0.002	0.169
6	0.357	0.997	0.943	0.971	0.999	0.002	0.140

Таблица 2. Исходные значения весов $W_0 = [w_{ij}]$

	A1	A2	A3	A4	A5	A6	A7
A1	0	0.1	0	0	0	-0.3	0
A2	0	0	0.7	0	0	0	0
A3	0.6	0	0	0	0	0	0
A4	0.9	0	0	0	0	0	0
A5	0	0	0.9	0	0	0	0
A6	0	0	0	0	-0.9	0	0.8
A7	0	0	0	0.9	-0.9	0	0

влияний, заданных в таблице 2. Определим веса w_{ij} взаимных влияний концепций по информации таблицы 1.

Определим матрицу Лапласа из соотношения (1):

$$L_0 = \begin{pmatrix} -0.2 & -0.3 & -0.2 & -0.2 & 0 & 0.1 & 0 \\ 0.7 & 0.7 & 0 & 0.7 & 0.7 & 0.7 & 1 \\ 0 & 0.6 & 0.6 & 0.6 & 0.6 & 0.6 & 1 \\ 0 & 0.9 & 0.9 & 0.9 & 0.9 & 0.9 & 1 \\ 0.9 & 0.9 & 0 & 0.9 & 0.9 & 0.9 & 1 \\ -0.1 & -0.1 & -0.1 & -0.1 & 0.8 & 0 & -1 \\ 0 & 0 & 0 & -0.9 & 0.9 & 0 & 0 \end{pmatrix}.$$

Собственные значения $e_0 = \text{eig}(L_0) =$

$$= \begin{pmatrix} 2.7123 & 0.7474 + 0.5012i & 0.7474 - 0.5012i & -0.5862 + 0.4398i & -0.5862 - 0.4398i & -0.1346 & 0 \end{pmatrix}.$$

Определим коэффициенты q разложения $Q(t)$ для лапласиана L_0 по формуле (7):

$$q_0 = \begin{pmatrix} -9.4853 & 0.0196 + 0.0477i & 0.0031 + 0.0152i & -0.0012 - 0.0008i & -0.0002 & 0 & 0 \end{pmatrix}.$$

Таблица 3. Исходный временной ряд

t	A1	A2	A3	A4	A5	A6	A7
0	0.300	0.350	0.400	0.450	0.500	0.550	0.600
1	0.153	0.656	0.592	0.750	0.853	0.614	0.589
2	0.085	0.947	0.735	0.877	0.989	0.294	0.527
3	0.116	0.993	0.813	0.901	0.997	0.034	0.443
4	0.187	0.994	0.878	0.933	0.997	0.007	0.364
5	0.258	0.995	0.921	0.954	0.998	0.004	0.306
6	0.331	0.997	0.947	0.964	0.999	0.003	0.269

Таблица 4. Исходные значения весов $W_0 = [w_{ij}]$

	A1	A2	A3	A4	A5	A6	A7
A1	0	0.097	0	0	0	-0.312	0
A2	0	0	0.681	0	0	0	0
A3	0.580	0	0	0	0	0	0
A4	0.885	0	0	0	0	0	0
A5	0	0	0.889	0	0	0	0
A6	0	0	0	0	-0.901	0	0.819
A7	0	0	0	0.925	-0.892	0	0

Определим матрицу Лапласа из соотношения (1):

$$L_1 = \begin{pmatrix} -0.173 & -0.28 & -0.173 & -0.173 & -0.173 & 0.1071 & -0.173 \\ 0.704 & 0.704 & 0 & 0.704 & 0.7042 & 0.7042 & 0.704 \\ 0 & 0.6 & 0.6 & 0.6 & 0.5998 & 0.5998 & 0.6 \\ 0 & 0.894 & 0.894 & 0.894 & 0.8941 & 0.8941 & 0.894 \\ 0.913 & 0.913 & 0 & 0.913 & 0.9129 & 0.9129 & 0.913 \\ -0.055 & -0.055 & -0.055 & -0.055 & 0.8212 & -0.055 & -0.876 \\ -0.012 & -0.012 & -0.012 & -0.918 & 0.9055 & -0.012 & -0.012 \end{pmatrix}.$$

Собственные значения $e_1 = \text{eig}(L_1) =$

$$= \begin{pmatrix} 2.6911 & -0.5909 + 0.4714i & -0.5909 - 0.4714i & 0.6422 + 0.4553i \\ 0.6422 - 0.4553i & 0.0064 & 0 & 0 \end{pmatrix}.$$

Определим коэффициенты q разложения $Q(t)$ для лапласиана L_1 по формуле (7):

$$q_1 = \begin{pmatrix} -9.6948 & 0.0060 - 0.0224i & -0.0022 + 0.0052i & -0.0005 + 0.0004i & 0.0001 & 0 & 0 \end{pmatrix},$$

тогда

$$q_1 - q_0 =$$

$$= \begin{pmatrix} -0.2091 & -0.0135 - 0.0701i & -0.0052 - 0.0099i & 0.0007 + 0.0012i & 0.0003 & 0 & 0 \end{pmatrix}.$$

Решая задачу минимизации (9), найдём матрицу Махаланобиса: $W = \text{diag}(0.0036 \ 0.0081 \ 0.0324 \ 0.1089 \ 0.1156 \ 0 \ 0)$ и расстояние Махаланобиса между векторами коэффициентов q_1 и q_0 : $WM_1 = (q_1 - q_0) \cdot W \cdot (q_1 - q_0)^T = 0.000203$.

Таблица 5. Исходный временной ряд

t	A1	A2	A3	A4	A5	A6	A7
0	0.300	0.350	0.400	0.450	0.500	0.550	0.600
1	0.170	0.646	0.636	0.770	0.842	0.624	0.488
2	0.111	0.961	0.761	0.890	0.988	0.251	0.281
3	0.169	0.993	0.843	0.907	0.997	0.011	0.106
4	0.288	0.996	0.894	0.948	0.997	0.002	0.046
5	0.430	0.997	0.951	0.968	0.998	0.001	0.041
6	0.601	0.998	0.969	0.983	0.998	0.001	0.044

Таблица 6. Исходные значения весов $W_0 = [w_{ij}]$

	A1	A2	A3	A4	A5	A6	A7
A1	0	0.151	0	0	0	-0.308	0
A2	0	0	0.727	0	0	0	0
A3	0.619	0	0	0	0	0	0
A4	0.841	0	0	0	0	0	0
A5	0	0	0.801	0	0	0	0
A6	0	0	0	0	-0.890	0	0.827
A7	0	0	0	0.860	-0.991	0	0

Определим матрицу Лапласа из соотношения (1):

$$L_2 = \begin{pmatrix} -0.157 & -0.308 & -0.157 & -0.157 & -0.157 & 0.1506 & -0.157 \\ 0.727 & 0.727 & 0 & 0.727 & 0.7266 & 0.7266 & 0.727 \\ 0 & 0.619 & 0.619 & 0.619 & 0.619 & 0.619 & 0.619 \\ 0 & 0.841 & 0.841 & 0.841 & 0.841 & 0.841 & 0.841 \\ 0.801 & 0.801 & 0 & 0.801 & 0.801 & 0.801 & 0.801 \\ -0.063 & -0.063 & -0.063 & -0.063 & 0.8268 & -0.063 & -0.89 \\ -0.131 & -0.131 & -0.131 & -0.991 & 0.8595 & -0.131 & -0.131 \end{pmatrix}.$$

Собственные значения $e_2 = \text{eig}(L_2) =$

$$= \begin{pmatrix} 2.7151 & -0.5839 + 0.4793i & -0.5839 - 0.4793i & 0.6149 + 0.4594i \\ & 0.6149 - 0.4594i & 0.0229 & 0 \end{pmatrix}.$$

Определим коэффициенты q разложения $Q(t)$ для лапласиана L_2 по формуле (7):

$$q_2 = \begin{pmatrix} -8.1360 & 0.0063 - 0.0282i & -0.0032 + 0.0065i & -0.0006 + 0.0005i & 0.0001 & 0 & 0 \end{pmatrix},$$

тогда

$$q_2 - q_0 = \begin{pmatrix} 1.3493 & -0.0133 - 0.0759i & -0.0063 - 0.0086i & 0.0005 + 0.0012i & 0.0003 & 0 & 0 \end{pmatrix}.$$

При матрице Махаланобиса: $W = \text{diag}(0.0036 \ 0.0081 \ 0.0324 \ 0.1089 \ 0.1156 \ 0 \ 0)$ найдём расстояние Махаланобиса между векторами коэффициентов q_2 и q_0 : $WM_2 = (q_2 - q_0) \cdot W \cdot (q_2 - q_0)^T = 0.0066$. Так как $WM_2/WM_1 \approx 32.5$, расстояние Махаланобиса WM_2 больше расстояния Махаланобиса WM_1 .

Заключение

В статье рассмотрена возможность сравнения временных рядов совокупности концепций. Для сравнения временных рядов предлагается формировать взвешенные ориентированные графы взаимных влияний концепций временных рядов друг на друга: вершинам графа соответствуют концепции, а взвешенным дугам значения величин влияния одной концепции на другую. Рассмотрено использование инвариантов, определяемых из теплового ядра уравнения теплопередачи, построенного на основе матрицы Лапласа графов, для характеристики графов. Применение метода сравнения временных рядов основе построения и сравнения графов позволит идентифицировать штатное и нештатное функционирование динамических объектов, систем управления, приборного состава и т.д. Предложен алгоритм минимизации расстояния Махаланобиса между вершинами графа, основанный на обучении матрицы Махаланобиса.

Благодарности

Работа выполнена в рамках федерального проекта «Экономика данных и цифровая трансформация государства»: «Кадры для цифровой трансформации», грант на обеспечение обучения студентов по образовательным программам высшего образования для топ-специалистов в сфере информационных технологий на базе АНО «Аналитический центр при Правительстве Российской Федерации» ИГК по договору ИТ: 000000Ц830325P2M0002.

Литература

1. Chung F.R.K. Spectral graph theory. American Mathematical Soc., 1997. V. 92.
2. Gilkey P.B. Invariance theory: the heat equation and the Atiyah-Singer index theorem. CRC press, 2018.
3. Xiao B., Hancock E.R., Wilson R.C. Graph characteristics from the heat kernel trace // Pattern Recognition. 2009. V. 42, No. 11. P. 2589–2606.
4. McDonald P., Meyers R. Isospectral polygons, planar graphs and heat content // Proceedings of the American mathematical society. 2003. V. 131, No. 11. P. 3589–3599.
5. Deza M.M., Deza E. Encyclopedia of Distances. Springer: Berlin-Heidelberg, 2016.
6. Mahalanobis P.C. On the generalised distance in statistics // Sankhya, A 80. 1936. P. 1–7.
7. Lu W. et al. The modeling and prediction of time series based on synergy of high-order fuzzy cognitive map and fuzzy c-means clustering // Knowledge-Based Systems. 2014. V. 70. P. 242–255.
8. Hagiwara M. Extended fuzzy cognitive maps // IEEE Transactions on Electronics, Information and Systems. 1994. V. 114, No. 3. P. 367–372.

COMPARISON OF TIME SERIES BASED ON THE ANALYSIS OF CORRESPONDING GRAPHS

S.N. Chukanov¹

Dr.Sc. (Techn.), Professor, Senior Scientist Researcher, e-mail: ch_sn@mail.ru

I.S. Chukanov²

master's student, e-mail: chukanov022@gmail.com

S.V. Leykhter³

Senior Lecturer, e-mail: leykhter@mail.ru

¹Sobolev Institute of Mathematics, Omsk branch, Russia

²Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia

³Omsk State University F.M. Dostoevsky

Abstract. This paper examines the comparison of time series of a set of concepts based on the formation of weighted directed graphs of the mutual influences of time series concepts on each other. The graph vertices correspond to the concepts, and the weighted arcs correspond to the values of the influence of one concept on another. The graph structure can be

characterized using the spectrum of the eigenvalues of the graph's Laplace matrix, which is related to the heat equation. For this purpose, a thermal kernel matrix is formed; the time derivative of the thermal kernel matrix is determined by the Laplace matrix. The coefficients of the Taylor series expansion of the thermal content of the thermal kernel can be used to represent the graph structure. The paper presents relations for determining the Mahalanobis distance between feature vectors thermal kernel invariants – which can be used for pattern recognition. The relevance of the problem is to increase the uniqueness of identification in time series recognition by taking into account both the spectrum of the eigenvalues of the graph's Laplace matrix and its eigenvectors. The aim of this study is to explore the feasibility of using the power series coefficients of the heat content expansion of the constructed graph as feature vectors—characteristics of the graph's properties. The novelty of the methods presented in this paper lies in the fact that, unlike the spectral embedding method, it explores the possibility of comparing graphs (images) based on the Mahalanobis distance between invariants constructed from the heat content expansion coefficients. An algorithm for minimizing the Mahalanobis distance between graph vertices is proposed, based on learning the Mahalanobis matrix. By comparing graphs corresponding to time series of a set of concepts, it is possible to compare time series.

Keywords: heat kernel on a graph, Mahalanobis distance, time series comparison.

Дата поступления в редакцию: 05.01.2026