Math-Net.Ru

Общероссийский математический портал

А. Э. Ерёменко, Доказательство Г. Шталя гипотезы
БМВ, *Матем. сб.*, 2015, том 206, номер 1, 97–102

DOI: <https://doi.org/10.4213/sm8294>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 34.229.45.116

26 апреля 2023 г., 14:05:19

УДК 517.53

А. Э. Ерёменко

Доказательство Г. Шталя гипотезы БМВ

Приведена упрощенная версия доказательства Г. Шталя гипотезы Бессиса–Муссы–Виллани о следах матриц $A + tB$, где A и B эрмитовы.

Библиография: 5 названий.

Ключевые слова: эрмитова матрица, теория возмущений, след, риманова поверхность.

DOI: 10.4213/sm8294

В статье приводится упрощенная версия доказательства Г. Шталя гипотезы Бессиса–Муссы–Виллани (БМВ); см. [1]. Доказательство использует основную идею Шталя, упрощение касается технических деталей.

ТЕОРЕМА. Пусть A и B – эрмитовы матрицы $n \times n$. Тогда функция

$$f(t) = \text{Tr} \exp(A - tB) \quad (1)$$

имеет представление

$$f(t) = \int_{b_1}^{b_n} e^{-ts} d\mu(s), \quad (2)$$

где μ – неотрицательная мера, а b_1 и b_n – наименьшее и наибольшее собственные значения матрицы B .

Если матрица B неотрицательно определенная, получаем $(-1)^n f^{(n)}(t) \geq 0$, $t \in \mathbb{R}$. Такие функции называются абсолютно монотонными. Теорема была высказана в качестве гипотезы в [2]. Известны две эквивалентные переформулировки (см. [3]) для неотрицательно определенных матриц B :

- а) коэффициенты многочлена $t \mapsto \text{Tr}(A + Bt)^p$, $p \in \mathbb{N}$, неотрицательны;
- б) функция $t \mapsto \text{Tr}(A + tB)^{-p}$ абсолютно монотонна при любом $p \geq 0$.

До работы Г. Шталя эта теорема была известна для случая матриц 2×2 . Обзор предшествующих попыток доказать эту теорему содержится в [4]. Доказательство Г. Шталя, которое излагается в нашей статье, вполне элементарно; все используемые средства существовали в середине 19 века.

Не умаляя общности, можно считать что B диагональна с собственными значениями $b_n > b_{n-1} > \dots > b_1 > 0$. Это достигается одновременным сопряжением A и B , прибавлением скаляра к B и аппроксимацией полученной матрицы B матрицей с простыми собственными значениями.

Собственные значения λ матрицы $A - tB$ определяются из уравнения

$$\det(\lambda I - A + tB) = 0.$$

Работа выполнена при поддержке Национального научного фонда, США (грант DMS-1361836).

© А. Э. ЕРЁМЕНКО, 2015

Этот определитель – многочлен от двух переменных t и λ . Вынося t за знак определителя и обозначая $y = \lambda/t$, $x = 1/t$, получаем уравнение вида

$$0 = \det(yI + B - xA) = \prod_{j=1}^n (y + b_j - xa_{j,j}) + O(x^2),$$

где $O(x^2)$ обозначает многочлен, делящийся на x^2 .

Отсюда следует, что существует n голоморфных ветвей многозначной функции $\lambda(t)$ в окрестности бесконечности, которые удовлетворяют соотношению

$$\lambda_j(t) = -b_j t + a_{j,j} + O\left(\frac{1}{t}\right), \quad t \rightarrow \infty, \quad (3)$$

причем все λ_j вещественны на вещественной оси. Кроме того, по теореме Реллиха (см. [5; теорема 12.4]) каждая из этих ветвей допускает аналитическое продолжение в некоторую область, содержащую вещественную ось. Алгебраическая функция $\lambda(t)$ определена на римановой поверхности S , n -листно накрывающей сферу Римана. Эта риманова поверхность не обязательно связна. Она имеет n неразветвленных листов над некоторой областью, которая содержит вещественную ось и окрестность бесконечности.

Рассмотрим частный случай теоремы, когда A тоже диагональна; тогда члена $O(1/t)$ в (3) нет, и мы получаем

$$f(t) = \sum_{j=1}^n e^{a_{j,j}} e^{-b_j t} = \int_0^\infty e^{-st} \sum_{j=1}^n e^{a_{j,j}} \delta_{b_j}(s) ds. \quad (4)$$

Так что в этом случае μ – дискретная мера с положительными атомами в собственных значениях матрицы B .

В общем случае дискретная компонента μ будет той же, но добавится непрерывная компонента – некоторая положительная функция на (b_1, b_n) . Г. Шталь нашел такое явное выражение для плотности меры μ .

ПРЕДЛОЖЕНИЕ. Мера

$$d\mu(s) = \left(\sum_{j=1}^n e^{a_{j,j}} \delta_{b_j}(s) + w(s) \right) ds, \quad (5)$$

где

$$w(s) = \frac{1}{2\pi i} \sum_{j: b_j < s} \int_C e^{\lambda_j(\zeta) + s\zeta} d\zeta, \quad (6)$$

удовлетворяет (1) и (2). Здесь C – любая окружность достаточно большого радиуса с центром в нуле, ориентированная против часовой стрелки.

Г. Шталь пишет, что догадаться до (5), (6) было трудно, и не объясняет, как он это сделал. Поэтому мы приведем эвристическое рассуждение, которое позволяет угадать эту формулу. Формула обращения преобразования Лапласа дает плотность в виде

$$\frac{1}{2\pi i} \int_L f(\zeta) e^{s\zeta} d\zeta,$$

где L – вертикальная прямая достаточно далеко справа. Когда $|\zeta|$ достаточно велико, выражение под интегралом есть

$$e^{s\zeta} f(\zeta) = \sum_{n=1}^n e^{\lambda_j(\zeta)+s\zeta}.$$

Так как $\lambda_j = -b_j\zeta + \dots$, где $b_j > s$ экспоненциально убывают в правой полуплоскости, то для этих слагаемых прямую L можно сдвигать вправо, и все эти слагаемые исчезают. Остальные слагаемые экспоненциально убывают влево, и для них контур можно изогнуть влево и превратить его в окружность C .

Конечно, можно строго обосновать это рассуждение, но как только правильная формула угадана, ее нетрудно проверить непосредственным вычислением, и мы воспроизведем вычисление Шталя.

ЛЕММА 1. *Для любого s выполняется*

$$\sum_{j=1}^n \int_C e^{\lambda_j(\zeta)+s\zeta} d\zeta = 0.$$

ДОКАЗАТЕЛЬСТВО. В самом деле, это интеграл от целой функции по замкнутому контуру.

Получается, что плотность w , определенная формулой (6), равна нулю при $s > b_n$. То, что она равна нулю при $s < b_1$, очевидно.

ДОКАЗАТЕЛЬСТВО ПРЕДЛОЖЕНИЯ. Вычисляем преобразование Лапласа от плотности w , определенной формулой (6):

$$\int_0^\infty e^{-st} w(s) ds = \sum_{k=1}^{n-1} \int_{b_k}^{b_{k+1}} e^{-ts} w(s) ds =: \sum_{k=1}^{n-1} I_k(t).$$

Фиксируем $t > 0$ и деформируем контур C в (6) так, чтобы положительный луч оказался вне C . Это возможно, потому что все λ_j голоморфны в некоторой области, содержащей вещественную ось и C . Число t теперь лежит вне деформированного контура C' . Согласно (6) имеем

$$I_k(t) = \int_{b_k}^{b_{k+1}} \sum_{j=1}^k \frac{1}{2\pi i} \int_{C'} e^{\lambda_j(\zeta)+s(\zeta-t)} d\zeta ds.$$

Меняя порядок интегрирования и порядок суммирования, получаем

$$\begin{aligned} \sum_{k=1}^{n-1} I_k(t) &= \sum_{j=1}^{n-1} \frac{1}{2\pi i} \int_{C'} e^{\lambda_j(\zeta)} \int_{b_j}^{b_n} e^{s(\zeta-t)} ds d\zeta \\ &= \sum_{j=1}^{n-1} \frac{1}{2\pi i} \int_{C'} e^{\lambda_j(\zeta)} (e^{b_n(\zeta-t)} - e^{b_j(\zeta-t)}) \frac{d\zeta}{\zeta-t}. \end{aligned}$$

Последнее выражение преобразуем с помощью формулы Коши, используя то, что t лежит вне C' . Получаем

$$\sum_{j=1}^n \int_{C'} e^{\lambda_j(\zeta)} e^{b_n(\zeta-t)} \frac{d\zeta}{\zeta-t} = 0,$$

подобно утверждению леммы. Таким образом,

$$\sum_{k=1}^{n-1} I_k(t) = - \sum_{j=1}^n \frac{1}{2\pi i} \int_{C'} e^{\lambda_j(\zeta) + b_j(\zeta - t)} \frac{d\zeta}{\zeta - t}.$$

Используя (3), можно записать

$$\lambda_j(\zeta) = -b_j\zeta + a_{j,j} + r_j(\zeta),$$

где $r_j(\infty) = 0$, и, применяя опять формулу Коши, получаем для любого j

$$\begin{aligned} -\frac{1}{2\pi i} \int_{C'} e^{\lambda_j(\zeta) + b_j(\zeta - t)} \frac{d\zeta}{\zeta - t} &= -\frac{e^{-b_j t + a_{j,j}}}{2\pi i} \int_{C'} e^{r_j(\zeta)} \frac{d\zeta}{\zeta - t} \\ &= e^{-b_j t + a_{j,j}} (e^{r_j(t)} - 1) = e^{\lambda_j(t)} - e^{-b_j t + a_{j,j}}. \end{aligned}$$

Складывая эти выражения при $j = 1, \dots, n$ и сравнивая с формулой (5) и со вторым равенством (4), получаем предложение.

Остается доказать, что (6) неотрицательно при любом s . Зафиксируем s и k так, что $b_k < s < b_{k+1}$. Замечательная идея Г. Штала состоит в том, чтобы заменить контур интегрирования в (6) на удачно выбранный гомологичный контур, по которому интеграл будет неотрицательным просто потому, что подынтегральная функция неотрицательна.

Напомним, что S – (возможно, несвязная) риманова поверхность, разветвленно накрывающая риманову сферу переменной ζ . Будем обозначать переменную точку поверхности S через p , и пусть $\pi: S \rightarrow \mathbb{C}$ – проекция на сферу ζ . Тогда λ – мероморфная функция на S , у которой все полюсы простые и лежат над $\zeta = \infty$.

Из асимптотических выражений (3) следует, что существует $R > 0$ такое, что для всех $j \leq k$ функции

$$\lambda_j(\zeta) + s\zeta = (s - b_j)\zeta + \dots \quad (7)$$

голоморфны при $|\zeta| > R$, вещественны на вещественной оси и их производные строго положительны при $\zeta > R$ и при $\zeta < -R$, тогда как при $j > k$ их производные строго отрицательны. Увеличивая, если нужно, R , можно достичь того, что при $|\zeta| > R/2$ и $j \leq k$ выражение $\text{Im}(\lambda_j(\zeta) + s\zeta)$ будет одного знака с $\text{Im} \zeta$, а при $|\zeta| > R/2$ и $j > k$ будет противоположного знака с $\text{Im} \zeta$.

Риманова поверхность S имеет антиконформную инволюцию, индуцированную комплексным сопряжением. Множество неподвижных точек этой инволюции состоит из n кривых, π -прообразов вещественной оси. Эти кривые разбивают S на две половины S^+ и S^- , которые отображаются друг на друга инволюцией. Проекции этих половин – верхняя и нижняя полуплоскости.

Выбираем $C = \{\zeta : |\zeta| = R\}$ в (6), где R выбрано раньше.

Рассмотрим открытые множества

$$\begin{aligned} D^+ &:= \{p \in S : |\pi(p)| < R, \text{Im } \pi(p) > 0, \text{Im}(\lambda(p) + s\pi(p)) > 0\}, \\ D^- &:= \{p \in S : |\pi(p)| < R, \text{Im } \pi(p) < 0, \text{Im}(\lambda(p) + s\pi(p)) < 0\} \end{aligned}$$

и

$$D = \text{int}(\overline{D^+} \cup \overline{D^-}).$$

Множество $\{p \in S : |\pi(p)| = R\}$ состоит из n попарно не пересекающихся окружностей $C_j \subset S$, которые нумеруем соответственно ветвям λ_j в (7), так что $\lambda = \lambda_j$ на C_j . Согласно сказанному после (7) окружности C_j при $j \leq k$ принадлежат ∂D , тогда как при $j > k$ они не пересекаются с $\overline{D}$.

Пусть D_1 – компонента¹ множества D , граница которой содержит некоторые окружности C_j . Мы хотим доказать, что

$$\sum_{j: C_j \subset \partial D_1} \int_{C_j} e^{\lambda(p) + s\pi(p)} d\pi(p) > 0, \quad (8)$$

где окружности ориентированы против часовой стрелки, что согласуется с их ориентацией как частей границы ∂D . Складывая эти неравенства по всем компонентам множества D , мы получим приведенную выше теорему.

В самом деле, каждая окружность C_j при $j \leq k$ принадлежит границе в точности одной компоненты множества D , а окружности C_j при $j > k$ не принадлежат границе множества D .

Каждая компонента D_1 множества D – это риманова поверхность конечного типа, граница которой состоит из нескольких кривых, параметризованных окружностью.

Эта параметризация кусочно гладкая, однако не обязательно ни гладкая, ни инъективная. Мы называем эти кривые граничными кривыми компоненты D_1 . Наш выбор числа R гарантирует, что часть границы компоненты D_1 , которая проектируется в C , – это в точности та цепь, по которой идет интегрирование в (8). Рассмотрим оставшуюся часть границы ∂D_1 , которая проектируется в $|\zeta| < R$.

ЛЕММА 2. *Никакая граничная кривая множества D над $\{\zeta : |\zeta| < R\}$ не может проектироваться в открытую верхнюю или нижнюю полуплоскость.*

ДОКАЗАТЕЛЬСТВО. В самом деле, предположим, что γ – граничная кривая, проекция которой не пересекает вещественную ось. Ориентируем ее как часть границы множества D , т.е. так, чтобы D была слева. Предположим, не умаляя общности, что γ проектируется в верхнюю полуплоскость. Пусть $g(p) = \lambda(p) + s\pi(p)$. Поскольку $\text{Im } g > 0$ в D^+ и $\text{Im } g = 0$ на γ , мы заключаем, что нормальная производная функции $\text{Im } g$ имеет постоянный знак на γ . Тогда по условиям Коши–Римана производная функции $\text{Re } g$ вдоль γ тоже имеет постоянный знак, что невозможно, потому что γ – замкнутая кривая, а однозначная функция $\text{Re } g$ не может строго возрастать на замкнутой кривой.

Таким образом, любая граничная кривая компоненты D_1 пересекает вещественную ось. Пусть γ – граничная кривая компоненты D_1 , проектирующаяся в $\{\zeta : |\zeta| < R\}$. По лемме 2 γ отображается на себя инволюцией, так что она

¹Используя принцип максимума, можно доказать, что любая компонента множества D имеет на границе некоторые окружности C_j , но мы это не используем.

состоит из двух симметричных кусков: один кусок, γ^+ , проектируется в верхнюю полуплоскость, а другой кусок, γ^- , — в нижнюю. Во всех концах p кривых γ^+ и γ^- выполняется $\operatorname{Im} \pi(p) = 0$. Имеем

$$e^g = e^{\operatorname{Re} g + i \operatorname{Im} g} = e^{\operatorname{Re} g} (\cos(\operatorname{Im} g) + i \sin(\operatorname{Im} g)).$$

Так как $\operatorname{Im} g = 0$ на γ и $\operatorname{Re} g$ возрастает, заключаем, что $\varphi(t) := e^{g(\gamma(t))}$ — вещественная возрастающая функция натурального параметра t на γ^+ . Поэтому

$$\begin{aligned} \frac{1}{2\pi i} \int_{\gamma} e^{g(p)} d\pi(p) &= \frac{1}{2\pi i} \left(\int_{\gamma^+} + \int_{\gamma^-} \right) \varphi(t) (d\xi(t) + i d\eta(t)) \\ &= \frac{1}{\pi} \int_{\gamma^+} \varphi(t) d\eta(t) = -\frac{1}{\pi} \int_{\gamma^+} \eta(t) d\varphi(t) < 0, \end{aligned}$$

где мы проинтегрировали по частям, используя $\eta(t) = 0$ на концах γ^+ .

Поскольку по теореме Коши интеграл от голоморфной 1-формы по границе равен нулю,

$$\int_{\partial D_1} e^g d\pi = 0,$$

то вклад в интеграл от той части границы ∂D_1 , которая проектируется в $\{\zeta : |\zeta| < R\}$, равен вкладу от части границы над C с противоположным знаком. Это завершает доказательство (8) и приведенной теоремы.

Список литературы

- [1] H. R. Stahl, “Proof of the BMV conjecture”, *Acta Math.*, **211**:2 (2013), 255–290.
- [2] D. Bessis, P. Moussa, M. Villani, “Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics”, *J. Mathematical Phys.*, **16**:11 (1975), 2318–2385.
- [3] E. H. Lieb, R. Seiringer, “Equivalent forms of the Bessis–Moussa–Villani conjecture”, *J. Statist. Phys.*, **115**:1–2 (2004), 185–190.
- [4] P. Moussa, “On the representation of $\operatorname{Tr}(e^{A-\lambda B})$ as a Laplace transform”, *Rev. Math. Phys.*, **12**:4 (2000), 621–655.
- [5] М. Рид, Б. Саймон, *Методы современной математической физики*, т. 4: *Анализ операторов*, Мир, М., 1982, 430 с.; пер. с англ.: М. Reed, В. Simon, *Methods of mathematical physics*, v. IV: *Analysis of operators*, Academic Press, New York–London, 1978, xv+396 pp.

Александр Эммануилович Ерёменко
(Alexandre É. Eremenko)

Purdue University, West Lafayette, IN, USA

E-mail: eremenko@math.purdue.edu

Поступила в редакцию

28.10.2013 и 04.04.2014