ОРДЕНА ЛЕНИНА ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ им.М.В.КЕЛДЫША
АКАДЕМИИ НАУК СССР

Е.А.Морозова, Н.Н.Ченцов

МАРКОВСКИЕ ИНВАРИАНТЫ
ПАР ВЕРОЯТНОСТНЫХ ЗАКОНОВ

Москва 1986

МАРКОВСКИЕ ИНВАРИАНТЫ ПАР ВЕРОЯТНОСТНЫХ ЗАКОНОВ

Е.А.Морозова, Н.Н.Ченцов (аннотация)

В математической статистике широко используются естественные коэффициенты различия между распределениями вероятностей на одном и том же пространстве случайных исходов, такие как относительная энтропия, вариация разности, риманово расстояние Бхаттачарья - Фишера и т.п. Все они монотонны относительно категории марковских отображений. В работе развивается общий подход к изучению подобных количеств как в классической, так и в некоммутативной теории вероятностей. В частности, установлено, что вариация разности двух законов является минорантой (с точностью до множителя) в классе всех марковски монотонных метрик. Как следствие, доказывается, что когда отсутствует априорная количественная информация о вероятностном законе, классическая задача статистической точечной оценки по независимым наблюдениям является некорректной при любой марковски монотонной функции потерь.

MARKOV INVARIANTS OF COUPLES OF PROBABILITY MEASURES

E.A. Morozova, N.N. Cencov

(abstract)

In the mathematical statistics the distinction between probability distributions on the same elementary event space is usually measured by some natural coefficient of divergence such as the relative entropy, the variation of substraction, the Riemann distance of Bhattachargya-Fisher, and so on. All of them are monotone under the category of Markov maps. In the paper the common approach to study such quantities is developped both in classical and in noncommutative probability theory. In particular, the variation of two distributions substraction is showed to be (up to the norming factor) the minorant in the class of all markovly monotone metrics. As the consequence, we prove the classical problem of statistical point estimation to be incorrect, when there is no a priori quantative about the observing probability distribution.

1. Совокупность всех вероятностных мер $P(\cdot)$ на измеримом пространстве (Ω , $\mathcal{A}$) элементарных исходов ω является выпуклым подмножеством полуупорядоченного линейного пространства всех счётно-аддитивных зарядов R , т.е. мер ограниченной вариации на (Ω , $\mathcal{A}$). С полуупорядоченностью этого пространства согласована естественная норма - полная вариация заряда

$$|R| = R^{+}(\Omega) + R^{-}(\Omega) , \qquad (I.1)$$

где R^{+} и R^{-} соответственно положительная и отрицательная части R :

$$R^{+}(A) = \sup_{B \subseteq A} R(B) , \quad R^{-}(A) = -\inf_{B \subseteq A} R(B) \qquad (I.2)$$

Естественная норма превращает совокупность всех вероятностных мер на (Ω , $\mathcal{A}$) в метрическое пространство с метрикой $\rho(P,Q) = |P - Q|$. Однако, хотя эта метрика и встречается в ряде вопросов теории вероятностей, в статистических проблемах с этой характеристикой близости двух распределений вероятностей дело иметь почти не приходится. Другой известной характеристикой близости двух вероятностных законов является расстояние Бхаттачария

$$s(P,Q) = 2 \arccos \int_{\Omega} \sqrt{P(d\omega)} \sqrt{Q(d\omega)} . \qquad (I.3)$$

Оно является римановой метрикой и задается для дискретного Ω дифференциальной квадратичной формой

$$(\delta s)^2 = \int_{\Omega} \frac{[(\delta P)(d\omega)]^{2}}{P(d\omega)} . \qquad (I.4)$$

Сама по себе форма (I.4), называемая "информацией по Фишеру", играет важную роль при формулировке и доказательстве неравенства информации в теории статистических оценок параметра, но про расстояние (I.3) "в большом" этого сказать нельзя.

В формулировку основных закономерностей теории проверки статистических гипотез входит информационное уклонение Кульбака-Лейблера-Санова

$$\text{Э}(P:Q) = \int_{\Omega} \left[\ln P(d\omega) - \ln Q(d\omega)\right] P(d\omega) , \qquad (I.5)$$

называемое также относительной энтропией. Как мы показали [I] ,

[2] , эта величина является своеобразным несимметричным аналогом (половины) квадрата расстояния между распределениями P и Q , и для этого расстояния справедлив несимметричный аналог теоремы Пифагора. Существуют и другие, правда несколько менее интересные характеристики близости двух распределений, своего рода несимметричные меры отличия одного распределения от другого, см. [2],[3].

Важным свойством перечисленных характеристик близости является их инвариантность относительно отображений совокупности вероятностных мер, индуцируемых обратимыми измеримыми отображениями выборочного пространства (Ω , $\mathcal{A}$). Эту инвариантность нетрудно вывести из формул для ρ , s , $Э$. Несколько сложнее проверить, что при необратимых отображениях индуцированные распределения вероятности сближаются в смысле указанных "характеристик непохожести". Как показано нами в [2] , это свойство "монотонной инвариантности" указанных величин имеет место и относительно более широкой системы всех марковских отображений, т.е. аффиных отображений Π вида

$$\Pi: \; P \to P\Pi = Q \; ; \quad Q(\cdot) = \int\limits_{\Omega} P(d\omega)\, \Pi(\omega; \cdot)\,, \qquad (1.6)$$

где $\Pi(\omega\,;\,\cdot\,)$ – переходное распределение вероятностей. Марковские отображения образуют алгебраическую категорию, см. [4], [5], также [2] . Классом ее объектов можно считать класс всех измеримых пространств (Ω , $\mathcal{A}$), или, как будем считать мы, изоморфный ему класс совокупностей всех распределений вероятностей на каждом из (Ω , $\mathcal{A}$).

Содержательный смысл марковских отображений состоит в том, что они описывают статистические решающие правила в смысле Вальда [6] . Оказывается, что два семейства вероятностных мер $\{P_\theta^{(i)}, \theta \in \Theta\}$ соответственно на ($\Omega^{(i)}$, $\mathcal{A}^{(i)}$), $i = 1,2$, пара с общим параметрическим множеством Θ , эквивалентны в теории статистического вывода, когда они конгруентны в категорной геометрии, т.е. когда существуют такие марковские отображения $Ш^{12}$ и $Ш^{21}$, что

$$P_\theta^{(1)} Ш^{12} = P_\theta^{(2)}\,, \quad P_\theta^{(2)} Ш^{21} = P_\theta^{(1)}\,, \quad \forall\, \theta \in \Theta\,. \qquad (1.7)$$

<u>Определение I.</u> Числовую функцию f , заданную на всех квадратах объектов категории, назовём инвариантом (пар), если

$$(P_1, P_2) \sim (Q_1, Q_2) \Rightarrow f(P_1, P_2) = f(Q_1, Q_2)\,, \qquad (1.8)$$

для всех конгруентных пар, и монотонным инвариантом если всегда

$$f(P_1, P_2) \geq f(P_1 Ш, P_2 Ш). \quad (1.9)$$

Очевидно, всякий монотонный инвариант является инвариантом в смысле (1.8).

Еще одной важной операцией является прямое умножение измеримых пространств и отвечающее ему тензорное умножение совокупностей всех вероятностных мер на них. Это умножение функционально относительно марковской категории, см. [7]. Заметим здесь, что относительная энтропия (1.5) аддитивна относительно этого умножения, что выделяет ее, а также форму (1.4) среди других характеристик непохожести.

2. Случайные явления микромира не описываются схемами классической (как теперь говорят, – коммутативной) теории вероятностей, потому что логика квантовых событий не является аристотелевой. Наиболее удобным способом задания объекта некоммутативной теории оказалась следующая алгебраическая схема. Задается некоторая инъективная алгебра фон Неймана $\mathcal{L}$ ограниченных линейных операторов, действующих на гильбертовом пространстве $\mathcal{H}$, например, алгебра $\mathcal{L}(\mathcal{H})$ всех таких операторов. Алгебра $\mathcal{L}$ является некоммутативным (вообще говоря) аналогом классической (коммутативной) алгебры всех ограниченных измеримых функций на пространстве элементарных исходов. Эрмитовы элементы алгебры $\mathcal{L}$ называются (ограниченными) наблюдаемыми. Вероятностное состояние объекта задается неотрицательным нормированным нормальным (т.е. ультраслабо непрерывным, или, что то же, монотонно непрерывным) линейным функционалом Φ на $\mathcal{L}$, $\Phi: \mathcal{L}^H \to \mathbb{R}$ — аналогом математического ожидания по вероятностной мере, см. [8], [9]. Совокупность $\mathfrak{S}(\mathcal{L})$ всех состояний Φ объекта – выпуклое замкнутое множество в предсопряженном пространстве $\mathcal{L}_*$, $(\mathcal{L}_*)^* = \mathcal{L}$. Эрмитовы идемпотенты E алгебры $\mathcal{L}$ являются ортопроекторами на соответствующие (замкнутые) подпространства $\mathcal{E} = E(\mathcal{H})$ пространства $\mathcal{H}$. По аналогии с коммутативным случаем, где идемпотен-

ты суть индикаторы измеримых множеств, эти подпространства называют событиями (а также "да - нет" - экспериментами). Очевидно классическая схема является частным случаем некоммутативной, с коммутативной алгеброй $\mathfrak{L}$ ограниченных измеримых функций на пространстве исходов.

<u>Определение 2.</u> По аналогии с коммутативным случаем естественно назвать марковским всякое аффинное отображение Π :

$\mathfrak{S}(\mathfrak{L}_1) \to \mathfrak{S}(\mathfrak{L}_2)$.

По линейности Π продолжается до монотонно непрерывного линейного отображения предсопряженных пространств (зарядов), Π : $\mathfrak{L}_{1*} \to \mathfrak{L}_{2*}$. Сопряженное отображение Π^*: $\mathfrak{L}_2^H \to \mathfrak{L}_1^H$ является $\mathbb{R}$ - линейным, неотрицательным, нормальным (монотонно непрерывным) и нормированным. Оно однозначно продолжается до $\mathbb{C}$ - линейного отображения $\mathfrak{L}_2 \to \mathfrak{L}_1$. Для простоты мы не будем различать Π и Π^* ; условимся, что Π пишется справа от состояния и слева от наблюдаемой. Нетрудно усмотреть, что система всех марковских отображений всех совокупностей $\mathfrak{S}(\mathfrak{L})$ образует алгебраическую категорию в смысле Элленберга - Маклейна [10] . Во-первых, тождественное отображение каждого объекта, очевидно, марковское. Во-вторых, композиция (последовательное выполнение) двух марковских отображений снова есть марковское отображение. В-третьих, эта операция композиции ассоциативна, поскольку это всегда имеет место для отображений. Ниже мы будем называть ее широкой марковской категорией.

Алгебры $\mathfrak{L}$ и их преддвойственные пространства можно тензорно перемножать. Стайнспринг [11] заметил, что индуцируемое этим умножением тензорное произведение марковских отображений может не быть неотрицательным и, тем самым, немарковским. Он выделил подкласс вполне положительных отображений, замкнутый относительно операции тензорного умножения.

Определение 3. Отображение $\Pi : \mathfrak{L}_2 \to \mathfrak{L}_1$ называется вполне положительным [11] , если все отображения $\Pi \otimes id_r$ положительны, где id_r тождественное отображение $\mathfrak{L}(\mathcal{H}_r)$ в себя, $\dim \mathcal{H}_r = r$.

$$\Pi \otimes id_r \geqslant 0 \quad , \quad \forall r \in \mathbb{N} \tag{2.1}$$

Вполне положительные марковские отображения также образуют категорию, которую мы далее будем называть узкой марковской категорией. В частности, результат взаимодействия квантовой частицы с частицами (случайной) среды описывается вполне положительным марковским отображением. Далее, по известной теореме Неймарка [12] о продолжении любого разложения единицы до ортогонального, выводится [11] , что все марковские отображения в коммутативную алгебру фон Неймана или же отображения такой алгебры являются вполне положительными. Последние описывают акт (классического) измерения какой-либо числовой характеристики частицы [13] .

Таким образом, определение 1 перенесенное в некоммутативную теорию дает нам две геометрических эквивалентности семейств состояний — широкую и узкую. Кроме того, по схеме определения 1 можно ввести эквивалентность семейств относительно измерений. Нетрудно установить для этих эквивалентностей цепочку импликаций

$$\{\text{узкая}\} \Rightarrow \{\text{широкая}\} \Rightarrow \{\text{измерительная}\} \tag{2.2}$$

Справедливо следующее предложение о максимальных марковски конгруентных семействах.

Теорема 2.1. Если вполне положительные марковские отображения $Ш^{12}$ и $Ш^{21}$ устанавливают конгурентность семейств $\{\Phi^{(i)}_\theta, \theta \in \Theta\} \subset \mathfrak{S}(\mathfrak{L}_i)$, $i = 1, 2$, то

$$\mathfrak{S}(\mathfrak{L}_i')Ш^{ij} = \mathfrak{S}(\mathfrak{L}_j') \supset \{\Phi^{(j)}_\theta, \theta \in \Theta\}, \ i \neq j = 1, 2, \tag{2.3}$$

где $\mathfrak{L}_i' \subset \mathfrak{L}_i$ — некоторые инъективные подалгебры фон Неймана.

Сходные утверждения справедливы и для широкой марковской конгурентности. Доказательство немедленно вытекает из эргодической теоремы для марковских эндоморфизмов [14] , [15] , см. также [16] , [17] .

3. В этой работе для удобства мы ограничимся сепарабельными пространствами $\mathcal{H}$. Условие инъективности $\mathfrak{L}$ позволит ограничиться полными алгебрами $\mathfrak{L}(\mathcal{H})$, $\mathfrak{S}(\mathfrak{L}(\mathcal{H})) = \mathfrak{S}(\mathcal{H})$, и использовать матричные обозначения для описания и наблюдаемых, и состояний и марковских отображений [18] , [19] . Для этого следует лишь фиксировать какой-либо ортонормированный базис в $\mathcal{H}$. Тогда наблюдаемые A будут описываться эрмитовыми матрицами $({}^i_j a)$, состояния Φ — неотрицательными эрмитовыми матрицами (φ^i_j) со следом единица. Чтобы различать эти объекты в первом случае индексы будут писаться слева. Для описанимарковского отображения $\Pi : \mathfrak{S}(\mathcal{H}^{(1)}) \to \mathfrak{S}(\mathcal{H}^{(2)})$ надо указать базисы в $\mathcal{H}^{(1)}$ и $\mathcal{H}^{(2)}$. Эти начальный и конечный базисы можно выбрать различными, даже если $\mathcal{H}^{(1)} = \mathcal{H}^{(2)}$.

Лемма 3.1. Марковское отображение $Ш : \mathfrak{S}(\mathcal{H}^{(1)}) \to \mathfrak{S}(\mathcal{H}^{(2)})$ описывается четырехвалентным (унитарным) тензором $({}^i_j ш^k_\ell)$, со свойствами

$$ {}^i_j Ш^k_\ell = \overline{{}^j_i Ш^\ell_k}, \quad \forall\, i, j, k, \ell ; \tag{3.1} $$

$$ tr_{\mathcal{H}^{(2)}} Ш := \sum_k {}^i_j ш^k_k = {}^i_j \delta ; \tag{3.2} $$

$$ \sum_{i,j,k,\ell} \xi^i_1 \xi_j \, {}^i_j Ш^k_\ell \, {}^k\eta \, {}_\ell\eta \geq 0 , \tag{3.3} $$

$$ \forall (\xi^1, \dots, \xi^n, \dots) = (\bar\xi_1, \dots, \bar\xi_n, \dots), \; \forall ({}^1\eta, \dots, {}^m\eta, \dots) = ({}_1\bar\eta, \dots {}_m\bar\eta, \dots). $$

Здесь ${}^i_j \delta$ - символ Кронекера, черта означает комплексное сопряжение, и подразумевается суммирование по совпадающим верхнему и нижнему или левому и правому индексам. Смысл, в котором понимается суммирование в счётномерном случае, установлен в [19] .

Условие (3.1) обеспечивает сохранение эрмитовости, условие (3.3) – сохранение неотрицательности, условие (3.2) – сохранение величины следа. Покажем теперь, как упрощается теория Стайнспринга [11] в конечномерном случае.

Лемма 3.2. Матрица (${}^{i}_{j}Ш^{k}_{\ell}$) вполне положительного марковского отображения удовлетворяет условиям (3.1), (3.2) и условию

$$\sum_{i,j,k,\ell} {}^{k}\zeta^{i}\ {}_{\ell}\zeta_{j}\ {}^{i}_{j}Ш^{k}_{\ell} \geq 0\,, \tag{3.4}$$

$$\forall\, ({}^{1}\zeta^{1}, \ldots, {}^{n}\zeta^{m}, \ldots) = ({}_{1}\bar{\zeta}_{1}, \ldots, {}_{n}\bar{\zeta}_{m}, \ldots)\,,$$

более ограничительному, чем условие (3.3).

Лемма 3.3. Матрица (${}^{i}_{j}Ш^{k}_{\ell}$) вполне положительного марковского отображения допускает [20] разложение

$$ {}^{i}_{j}Ш^{k}_{\ell} = \sum_{s} {}^{i}Щ^{k}(s)\, {}_{j}Щ_{\ell}(s)\,, \tag{3.5}$$

$$ {}^{i}Щ^{k}(s) = \overline{{}_{i}Щ_{k}(s)}\,, \quad \forall\, i, k, s\,.$$

Следствие. Композиция вполне положительных марковских отображений вполне положительна.

Доказательство. Из условия (3.4) вытекает, что матрица (${}^{i}_{j}Ш^{k}_{\ell}$) может быть интерпретирована, как матрица неотрицательного Эрмитова оператора, действующего в пространстве $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$, откуда следует (3.5). Далее, любой неотрицательный ядерный Эрмитов оператор в пространстве $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$ также разлагается в сумму тензоров вида $b^{ip}\, b_{jq}$, а $Ш \otimes id_{2}$ – образ каждого такого тензора неотрицателен в силу

$$\sum b^{ip}\, b_{jq}\ {}^{i}Щ^{k}\ {}_{j}Щ_{\ell}\ {}^{p}\delta^{u}\ {}_{q}\delta_{v}\ {}^{ku}a\ {}_{\ell v}a =$$

$$= \left| \sum \; b^{ip} \; {}^{i}Ш^{k} \; {}^{kp}a \right|^{2} \geq 0 .$$

Наконец, условие (3.4) необходимо для вполне положительности, потому что неотрицательный Эрмитов оператор в $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(1)}$ с матрицей ($\delta^{ip} \; \delta_{jq}$) отображается $Ш \otimes id_{\tau}$ в оператор с матрицей $\left(f^{ku}_{lv} \right)$, где $f^{ku}_{lv} = {}^{u}_{v}Ш^{k}_{l}$ при всех значениях индексов. В заключение отметим, что разложение Крауса справедливо и в счётномерном случае.

<u>4.</u> Взвешенная разность двух состояний одного объекта принадлежит эрмитовой части предсопряженного пространства $\mathcal{L}^{H}_{*} = \mathrm{Lin}\, \mathfrak{S}(\mathcal{L})$. Она описывается ядерным оператором.

<u>Определение 5.</u> Пусть $\Gamma \in \mathcal{L}^{H}_{*}$. Его положительную $\Gamma^{(+)}$ и отрицательную $\Gamma^{(-)}$ части определим через спектральное разложение:

$$\Gamma = \sum \gamma_k E_k \; , \quad \Gamma^{(+)} = \sum_{\gamma > 0} \gamma_k E_k \; , \quad \Gamma^{(-)} = - \sum_{\gamma < 0} \gamma_k E_k \, . \qquad (4.1)$$

Величины

$$tr^{(+)} \Gamma = tr\, \Gamma^{(+)} = \sum_{\gamma > 0} \gamma_k \; , \; tr^{(-)} \Gamma = - \sum_{\gamma < 0} \gamma_k \; , \; |\Gamma| = \sum |\gamma_k| \qquad (4.2)$$

назовем соответственно положительной, отрицательной и полной вариацией Γ . Разложения $\Gamma = \Gamma^{(+)} - \Gamma^{(-)}$, и $\mathcal{H} = \mathcal{E}^{(+)} \oplus \mathcal{E}^{(-)}$, где $E^{(+)} = \sum_{\gamma > 0} E_k$, $E^{(-)} = \sum_{\gamma < 0} E_k$ назовем расщеплениями.

<u>Лемма 4.1.</u> Положительная и отрицательная вариации $\Gamma \in \mathcal{L}^{H}_{*}$ являются монотонными инвариантами в широкой марковской категории. Каждое $\Gamma \in \mathcal{L}^{H}_{*}$ конгруентно заряду Γ_{red} на двухатомном измеримом пространстве $\Omega_2 = \{ \omega_1 , \omega_2 \}$

$$\Gamma_{red}(\omega_1) = tr^{(+)} \Gamma \, , \quad \Gamma_{red}(\omega_2) = - tr^{(-)} \Gamma \, . \qquad (4.3)$$

Следствие. Положительная и отрицательная вариации $\Gamma \in \mathcal{L}_*^{H}$ образуют полную систему инвариантов Γ в обоих марковских категориях.

Доказательство. Выбрав за базисные в $\mathcal{H}^{(1)}$ и $\mathcal{H}^{(2)}$ собственные векторы Γ и $\Gamma Ш$, проводим построения, известные в коммутативной теории, см. [2], лемма 5.9.

Применяя эти результаты к линейной комбинации состояний $\Psi - z\Phi$, получаем два семейства

$$\mu^{(\pm)}(z) = tr(\Psi - z\Phi)^{(\pm)}, \quad \mu^{(+)}(z) - \mu^{(-)}(z) = 1 - z, \ \forall z \in \mathbb{R}, \quad (4.4)$$

монотонных инвариантов пары (Φ, Ψ). Любое из них задаёт полную систему инвариантов, когда Φ и Ψ коммутируют как операторы, см. [2].

В общем случае вопрос о легко обозримой полной системе инвариантов и возможных аналогах формул (4.5) и (4.6) пока не решён. Соответствующее наше утверждение в [21] является ошибочным, так как коммутируемость состояний есть инвариантное свойство в узкой марковской категории.

5. Пространства с метрикой должны удовлетворять аксиомам 1^0 $\rho(x,y) \geq 0$, 2^0 $\rho(x,y)=0 \Rightarrow y=x$, 3^0 $\rho(x,y)=\rho(y,x)$, 4^0 $\rho(x,z) \leq \rho(x,y)+\rho(y,z)$. Когда выполнены только аксиомы 1^0, 3^0, 4^0, говорят о псевдометрике. Мы рассмотрим более экзотический случай несимметричной метрики, когда не выполнена аксиома 3^0(но аксиома 4^0 выполнена именно в указанной форме).

Теорема 5.1. Если на всех объектах $\mathfrak{S}(\mathcal{L})$ задана метрика $\rho(\Phi, \Psi)$, монотонно инвариантная в узкой марковской категории, то

$$\rho(\Phi, \Psi) \geq \frac{1}{8}\, \rho(R_{1/2}, R_{1/4}) \cdot |\Phi - \Psi|, \qquad (5.1)$$

где $|\Phi-\Psi|$ — полная вариация (следовая норма) разности $\Phi-\Psi$, R_θ - распределение вероятности на алгебре с двумя атомами,

$$R_\theta : \quad \{ R(\omega_1;\theta)=\theta \quad , \quad R(\omega_2;\theta)=1-\theta \} . \tag{5.2}$$

Замечание. Если метрика ρ не симметрична, то

$$\rho(\Phi,\Psi) \geq \frac{1}{8}|\Phi-\Psi| \cdot \min\{\rho(R_{1/2},R_{1/4}),\rho(R_{1/4},R_{1/2})\}.$$

Доказательство. Пусть $\mathcal{H}=\mathcal{E}^{(+)}\oplus\mathcal{E}^{(-)}$ - расщепление для ядерного оператора $\Phi-\Psi$. Сужение состояний на двухатомную булеву алгебру с атомами $\mathcal{E}^{(+)}$ и $\mathcal{E}^{(-)}$ леммы 4.1 - вполне положительное марковское отображение. Отсюда $\rho(\Phi,\Psi)\geq$ $\geq\rho(\Phi_{red},\Psi_{red})$ и $|\Phi-\Psi|=|\Phi_{red},\Psi_{red}|$.

Следовательно, достаточно доказать теорему для распределений вероятностей на двухатомной булевой алгебре. Каждое распределение вероятностей R_θ на ней описывается (5.2) вероятностью первого атома $R(\omega_1,\theta)=\theta$, $0\leq\theta\leq 1$. Марковский эндоморфизм совокупности $\{R_\theta, 0\leq\theta\leq 1\}$ описывается аффинным отображением отрезка $\{\theta : 0\leq\theta\leq 1\}$ в себя. При этом

$$|R_{\theta'}-R_{\theta''}|=2|\theta'-\theta''| . \tag{5.3}$$

Лемма 5.2. Если $0\leq\theta_1'\leq\theta_1''<\theta_2''\leq\theta_2'\leq 1$, то марковский эндоморфизм Π совокупности $\{R_\theta\}$ с матрицей

$$\begin{pmatrix} \dfrac{(1-\theta_1')(1-\theta_2'')-(1-\theta_2')(1-\theta_1'')}{\theta_2'-\theta_1'} & \dfrac{(1-\theta_1')\theta_2''-(1-\theta_2')\theta_1''}{\theta_2'-\theta_1'} \\[2ex] \dfrac{\theta_2'(1-\theta_1'')-\theta_1'(1-\theta_2'')}{\theta_2'-\theta_1'} & \dfrac{\theta_2'\theta_1''-\theta_1'\theta_2''}{\theta_2'-\theta_1'} \end{pmatrix} \tag{5.4}$$

отображает $R_{\theta_1'}$ в $R_{\theta_1''}$, $R_{\theta_2'}$ в $R_{\theta_2''}$.

Следствие. Если $f(\Phi,\Psi)$ - монотонный инвариант узкой марковской категории, то

$f(R_{\theta_1'}, R_{\theta_2'}) \geqslant f(R_{\theta_1''}, R_{\theta_2''})$ в условиях леммы.

<u>Лемма 5.3.</u> Пусть $0 \le \theta_0 \le 1$, $0 \le q \le 1$, $\theta_k = \theta_0 q^k$. Тогда марковские эндоморфизмы совокупности $\{R_\theta\}$ с матрицами

$$Ш^n = \begin{pmatrix} q^n & 1-q^n \\ 0 & 1 \end{pmatrix} \tag{5.5}$$

отображают по правилу $R_{\theta_k} Ш^n = R_{\theta_{k+n}}$. При этом

$$|R_{\theta_{k-1}} - R_{\theta_k}| = q^{k-1} |R_{\theta_0} - R_{\theta_1}| \tag{5.6}$$

Обе леммы доказываются простыми вычислениями. Перейдем к доказательству теоремы. Пусть $\Phi_{red} = (p_1, 1-p_1)$, $\Psi_{red} = (p_2, 1-p_2)$, $\theta_0 = \max\{p_1, p_2, 1-p_1, 1-p_2\}$. Не ограничивая общности, можно считать $\theta = p_1$. В противном случае перенумеруем исходы (что описывается обратимым марковским отображением распределений) или переставим аргументы. Заметим, сверх того, что при $p_1 = p_2$ неравенство тривиально выполнено. Поэтому мы можем считать, что $\theta_0 = p_1 > p_2$ и $\theta_0 = p_1 > 1/2$. Обозначим $q = p_2 \cdot p_1^{-1} < 1$ и построим точки $\theta_k = \theta_0 \cdot q^k$, $k = 1, \dots, \nu$, где $\theta_{\nu-1} > \frac{1}{4}$, $\theta_\nu < 1/4$, так что $q^k = \theta_k \cdot \theta_0^{-1} > 1/4$, $k = 0, 1, \dots, \nu-1$. По монотонности ρ, неравенству треугольника, следствию из леммы 5.2 и оценке (5.6)

$$\nu \cdot \rho(\Phi_{red}, \Psi_{red}) \geqslant \rho(R_{\theta_0}, R_{\theta_1}) + \dots + \rho(R_{\theta_{\nu-1}}, R_{\theta_\nu}) \geqslant$$

$$\geqslant \rho(R_{\theta_0}, R_{\theta_\nu}) \geqslant \rho(R_{1/2}, R_{1/4});$$

$$\nu \cdot |\Phi_{red} - \Psi_{red}| = 2|\theta_0 - \theta_1| + \dots + 2q^{-\nu+1} |\theta_{\nu-1} - \theta_\nu| \leqslant$$

$$\leqslant 8 |\theta_0 - \theta_\nu| \leqslant 8 .$$

Поделив первое из неравенств на второе, получаем (5.1) для двухатомных распределений вероятностей, что и требовалось. Для несимметричных расстояний может быть $p_2 > p_1$. Тогда надо строить точки $\theta_k = \theta_0 \cdot q^k$, где $\theta_0 = p_2$, $q = p_1 \ p_2^{-1}$.

6. Изучим строение естественных функций потерь $L(\Phi, \Psi)$ в задаче оценки состояния. Мы предполагаем, что функция L задана на всех квадратах объектов $\mathfrak{S}(\mathfrak{X})$, неотрицательна:

$$L(\Phi, \Phi) = 0 \ ; \ \Phi \neq \Psi \ \Rightarrow \ L(\Phi, \Psi) > 0, \qquad (6.1)$$

и монотонна в узкой категории:

$$L(\Phi\Pi, \Psi\Pi) \leq L(\Phi, \Psi). \qquad (6.2)$$

Мы допускаем, что L может принимать значение $+\infty$.

Рассмотрим сперва значения $\ell(x, y) = L(R_x, R_y)$ на двухатомной булевой алгебре.

<u>Лемма 6.1.</u> Функция $\ell(x,y)$ на квадрате $0 \leq x, y \leq 1$ тождественно равна нулю на диагонали $x = y$ и строго положительна вне неё. Функция

$$f(z) = \inf_{|x-y| \geq z} \ell(x, y) \qquad (6.3)$$

монотонна и строго положительна при $z > 0$, $f(0) = 0$.

Доказательство. Первые два утверждения следует из (6.1), а монотонность f следует из определения (6.3). Рассмотрим подмножество $Q(z) = \{(x,y) : |x-y| \geq z, \ 0 \leq x, y \leq 1\}$. Оно компактно. Поэтому, существует сходящаяся последовательность

$$(x_n, y_n) \to (x^*, y^*) \in Q(z), \quad \lim \ell(x_n, y_n) = f(z).$$

Пусть $x^* < y^*$. Возьмем любой отрезок $[x_0, y_0]$ такой, что $x^* < x_0 < y_0 < y^*$. Тогда $[x_0, y_0] \subset [x_n, y_n]$ при $n \geq n_0$. Отсюда $\ell(x_0, y_0) \leq \ell(x_n, y_n)$ при $n \geq n_0$, и $0 < \ell(x_0, y_0) \leq f(z)$. Если $x^* > y^*$, проведенное рассуждение надо применить к отрезкам вида $[y, x]$.

Теорема 6.2. Если заданная на квадратах объектов $\mathfrak{S}(\mathcal{L})$ функция L удовлетворяет условиям (6.1) и (6.2), то

$$L(\Phi,\Psi) \geqslant f(|\Phi-\Psi|), \tag{6.4}$$

где функция $f(z)$, определенная формулой (6.3), монотонна и строго положительна при $z>0$, $f(0)=0$.

Доказательство. По монотонности L и лемме 4.1, примененной к разности $(\Phi-\Psi)\in\mathcal{L}_*$, $L(\Phi,\Psi)\geqslant$ $\geqslant L(\Phi_{red},\Psi_{red})\geqslant f(|\Phi_{red}-\Psi_{red}|)=f(|\Phi-\Psi|)$, где среднее неравенство вытекает из леммы 6.1.

Если монотонная функция $F(\Phi,\Psi)$ вместо (6.1) удовлетворяет условию $F(\Phi,\Psi)\neq F(\Phi,\Phi)$ при $\Phi\neq\Psi$, то $F(\Phi,\Psi)\geqslant c+f(|\Phi-\Psi|)$, где $c=F(\Phi,\Phi)=F(\Psi,\Psi)$, а функция f построена по (6.3) для $L(\Phi,\Psi)=F(\Phi,\Psi)-c$.

7. Инвариантные римановы метрики особенно интересны тем, что определяющие их эквивариантные дифференциальные квадратичные формы позволяют выписывать информационные неравенства, см. [13].

Фиксируем какой-либо собственный базис ядерного оператора состояния Φ, $\mathcal{H}=\oplus\,\mathcal{L}^{(k)}$. В нем Φ будет описываться диагональной матрицей $diag\langle p_1,p_2,\dots\rangle$, где $p_i=\varphi_i^i$ - собственные числа. В этом же базисе дифференциал $d\Phi$ будет описываться эрмитовой матрицей со следом нуль:

$$d\varphi_\ell^k=\xi_\ell^k+i\,\eta_\ell^k, \tag{7.1}$$

где ξ и η - вещественные и мнимые части дифференциалов, а символ d опущен для упрощения обозначений. Имеют место связи

$$\xi_\ell^k=\xi_k^\ell,\ \eta_\ell^k=-\eta_k^\ell,\ \forall k,\ell;\ \eta_k^k=0,\forall k;\ \sum_k\xi_k^k=0. \tag{7.2}$$

Таким образом, полную систему линейных вещественных координат в касательном к $\mathfrak{S}$ пространстве образуют переменные ξ_ℓ^k, η_ℓ^k

при $k<\ell$, а также все переменные ξ_k^k, кроме одной из них. Для упрощения удобно записать искомую квадратичную форму через все ξ_k^k, см. [5], с помощью коэффициентов u, v, w полярной билинейной формы:

$$Q_\Phi(d\Phi) =$$

$$= \sum u_{j\ell}^{ik}\,\xi_j^i\xi_\ell^k + 2\sum v_{j\ell}^{ik}\,\xi_j^i\eta_\ell^k + \sum w_{j\ell}^{ik}\,\eta_j^i\eta_\ell^k \,. \tag{7.3}$$

Простейшие вполне положительные марковские отображения действуют по формуле

$$\Phi Ш_S = S\Phi S^*\,; \quad \left({}_j^i ш_\ell^k\right) = \left({}^i s^k{}_j s_\ell\right), \tag{7.4}$$

где S – унитарный оператор. Отображение $Ш_S$, очевидно, обратимо, и $\Phi \sim \Phi Ш_S$, $d\Phi \sim d\Phi Ш_S$.

<u>Лемма 7.1.</u> Если дифференциальная квадратичная форма (7.3) унитарно инвариантна, $Q_\Phi(d\Phi) = Q_{U\Phi U^*}(dU\Phi U^*)$, то она тождественно равна

$$\sum u_{jj}^{jj}\,\xi_j^j\xi_j^j + \sum\left(u_{jk}^{jk}+u_{kj}^{kj}\right)\xi_j^j\,\xi_k^k + \tag{7.5}$$

$$+ 2\sum u_{kk}^{jj}\left[(\xi_k^j)^2 + (\eta_k^j)^2\right].$$

Доказательство. Пусть симметрия S есть отображение первой оси $\mathcal{L}^{(1)}$. Оно меняет знак у элементов первой строки и первого столбца, кроме диагонального. Так как $(\Phi, d\Phi)\sim(\Phi Ш_S, d\Phi Ш_S)$, то значение формы Q сохраняется, и оно тождественно равно полусумме исходного Q и выражения, отличающегося знаком минус при членах, в которые индекс 1 входит один или три раза. В полусумме эти члены взаимно уничтожаются, а остальные сохраняются. Проделаем эту операцию последовательно при всех $k = 1, 2, \ldots$ В результате придем к форме

$$Q_\Phi(d\Phi) = \sum u^{jj}_{jj}(\xi^j_j)^2 + \sum (u^{jk}_{jk} + u^{kj}_{kj})\,\xi^j_j\,\xi^k_k + \tag{7.6}$$

$$+\sum u^{jj}_{kk}(\xi^j_k)^2 + 2\sum v^{jj}_{kk}\,\xi^j_k\,\eta^j_k + \sum w^{jj}_{kk}(\eta^j_k)^2,$$

где все двойные суммы взяты по парам $j < k$.

Теперь применим к аргументам Φ и $d\Phi$ операцию поворота первой "оси" в себе, индуцированного умножением на мнимую единицу. При этом $\xi^1_k \to \eta^1_k$, $\eta^1_k \to -\xi^1_k$. Снова усредняя полученные результаты, умножаем члены при $\xi^1_k\,\eta^1_k$ и объединяем слагаемые $(\xi^1_k)^2 + (\eta^1_k)^2$. Если в исходной форме положить все $\eta^1_k = 0$, то из Π – инвариантности формы Q получаем $u^{11}_{kk} = w^{11}_{kk}$. Проделав повороты при всех $j = 1, 2, \ldots$, приходим от (7.6) к (7.5).

<u>Теорема 7.2.</u> Если дифференциальная квадратичная форма (7.3) инвариантна в узкой марковской категории, то она тождественно равна

$$Q_\Phi(d\Phi) = c\sum_k (p_k)^{-1}(\xi^k_k)^2 + 2\sum_{j<k} c(p_j, p_k)\,|d\varphi^j_k|^2, \tag{7.7}$$

где $c(x, y)$ – функция в области $0 \leq x, y \leq x + y \leq 1$, причём $c(x, x) = c\,x^{-1}$, $c(x, y) = c(y, x)$.

<u>Доказательство.</u> В коммутативной теории инвариантная риманова метрика единственна с точностью до множителя. Отсюда по (I.10) $u^{jk}_{jk} = 0$ при $j \neq k$, $u^{kk}_{kk} = c/p_k$, и первые два слагаемые (7.5) дают первое слагаемое (7.7). Построим теперь марковское отображение $Ш$ из $\mathfrak{S}(\mathcal{L})$ на $\mathfrak{S}_{2,1} = \mathfrak{S}(\mathcal{L}(\mathcal{L}^{(j)} \oplus \oplus \mathcal{L}^{k}) \oplus \mathbb{C})$ по формуле $\Psi Ш = \Gamma \oplus \gamma$,

$$\Gamma = (E_j + E_k)\Psi(E_j + E_k), \quad \gamma = \operatorname{tr}\Psi - \psi^j_j - \psi^k_k, \tag{7.8}$$

и отображение $Ж : \mathfrak{S}_{2,1} \to \mathfrak{S}(\mathfrak{L})$

$$(\Gamma \oplus \gamma) Ж = \Gamma \oplus \gamma (H - E_j - E_k) \Phi (H - E_j - E_k) . \qquad (7.9)$$

Легко проверить по критерию (3.4) их вполне положительность. На операторе Φ и дифференциале $d\Phi = \xi^j_k$ отображения $Ш$ и $Ж$ взаимно обратны, $\gamma(\Phi) = 1 - p_j - p_k$. Отсюда $u^{jj}_{kk}(\Phi) = u^{jj}_{kk}(\Phi Ш) = u^{jj}_{kk}(p_j, p_k)$. Поскольку перенумерация осей $\mathcal{L}^{(k)}$ есть унитарное преобразование, то $u^{jj}_{kk}(x, y) = c(x, y)$.

Если $p_j = p_k$, то любой поворот плоскости $\mathcal{L}^{(j)} \oplus \mathcal{L}^{(k)}$ оставляет Φ инвариантным. Поворот Π на $\pi/4$ переводит $d\Phi = (\xi^j_k, \eta^j_k)$ в $d\Phi\Pi = (\gamma^j_j, \gamma^k_k)$, $\gamma^j_j = |\xi^j_k + i\eta^j_k|$, $\gamma^k_k = -\gamma^j_j$. Что дает $c(p,p) = c/p$.

<u>Лемма 7.3.</u> Если инвариантная дифференциальная квадратичная форма Q_Φ слабо непрерывна по Φ, то функция $c(x,y)$ в ее разложении (7.7) положительно однородна в минус первой степени

$$c(x,y) = \lambda \cdot c(\lambda x, \lambda y), \quad \forall \lambda : 0 < \lambda \leq (x+y)^{-1} \qquad (7.10)$$

Доказательство. Рассмотрим марковское отображение размножения $\Pi : \mathfrak{S}(\mathfrak{L}) \to \mathfrak{S}(\mathfrak{L} \otimes \mathfrak{L}(\mathcal{H}_n))$, действующее по формуле

$$\Psi \to \Psi \otimes \tfrac{1}{n} \mathbb{1}_n . \qquad (7.11)$$

Оно обратимо на образе $\mathfrak{S}(\mathfrak{L})\Pi$. По конгруентности $(\Phi, d\Phi)$ и его образа

$$c(x,y) = n \cdot c(xn^{-1}, yn^{-1}) = \nu \cdot c(x\nu^{-1}, y\nu^{-1})$$

Обозначая $x' = xn^{-1}$, $y' = yn^{-1}$, приходим к (7.10) с рациональным λ. Справедливость в общем случае следует из непрерывности.

<u>Теорема 7.4.</u> Всякая инвариантная дифференциальная квадра-

тичная форма со свойством (7.10), в частности слабо непрерывная, удовлетворяет принципу сложения информации

$$Q_{\Phi\otimes\Psi}(d[\Phi\otimes\Psi]) = Q_{\Phi}(d\Phi) + Q_{\Psi}(d\Psi). \qquad (7.12)$$

Доказательство. Выберем собственные базисы операторов $\Phi = \mathrm{diag}\langle p_1, p_2, \ldots\rangle$ и $\Psi = \mathrm{diag}\langle q_1, q_2, \ldots\rangle$. В тензорном произведении этих базисов только матрица дифференциала $d(\Phi\otimes\Psi) = \Phi\otimes(d\Psi) + (d\Phi)\otimes\Psi$ имеет много нулей, и

$$Q(d[\Phi\otimes\Psi]) = \sum_{i,k}(p_i q_k)^{-1}[p_i\, d\psi_k^k + q_k\, d\varphi_i^i]^2 +$$

$$+2\sum_{i;k<\ell} c(p_i q_k, p_i q_\ell)\,|p_i\, d\psi_\ell^k|^2 + 2\sum_{i<j;k} c(p_i q_k, p_j q_k)\,|q_k\, d\varphi_j^i|^2 \leq$$

$$\leq \sum_k (q_k)^{-1}|d\psi_k^k|^2 + \sum_i (p_i)^{-1}|d\varphi_i^i|^2 +$$

$$+2\sum_{k<\ell} c(q_k, q_\ell)\,|d\psi_\ell^k|^2 + 2\sum_{i<j} c(p_i, p_j)\,|d\varphi_j^i|^2,$$

где мы воспользовались правилом (7.10) и равенствами

$$\sum_i p_i = \sum_k q_k = 1, \quad \sum_i d\varphi_i^i = \sum_k d\psi_k^k = 0.$$

<u>8.</u> Если предположить, что риманова метрика с формой Q не только инвариантна, но и монотонна в категории, для коэффициента $c(x,y)$ формы (7.7) можно указать оценки.

<u>Лемма 8.1.</u> Если метрика монотонна, то

$$c(x,y) \geqslant \theta^2\cdot c(\theta x, \theta y), \quad \forall\,\theta : 0\leq\theta\leq 1. \qquad (8.1)$$

<u>Следствие.</u> В условиях леммы свойство (7.10) выполняется при всех λ без требования слабой непрерывности.

Доказательство. Зададим вполне положительный марковский эндоморфизм $Ш$ совокупности $\mathfrak{S}_{2,1} = \mathfrak{S}(\mathcal{L}(\mathcal{H}_2)\oplus\mathbb{C})$ состояний $\Psi = P\oplus q$, $q + \mathrm{tr}\,P = 1$, формулой

$$\Psi Ш = \theta P \oplus [q + (1-\theta)\, tr\, P] . \tag{8.2}$$

Его можно продолжить до эндоморфизма совокупности $\mathfrak{S}(\mathcal{H}_3) \supset \mathfrak{S}_{2,1}$, описываемого в базисе $\mathcal{H}_2 \oplus \mathbb{C} = \mathcal{H}_3$ четыре - матрицей третьего порядка

$$\left(\begin{array}{ccc:ccc:c}
\theta\ 0\ 0 & & & 0\ \theta\ 0 & & & \\
0\ 0\ 0 & & & 0\ 0\ 0 & & & 0 \\
0\ 0\ 1-\theta & & & 0\ 0\ 0 & & & \\
\hdashline
0\ 0\ 0 & & & 0\ 0\ 0 & & & \\
\theta\ 0\ 0 & & & 0\ \theta\ 0 & & & 0 \\
0\ 0\ 0 & & & 0\ 0\ 1-\theta & & & \\
\hdashline
 & & & & & & 0\ 0\ 0 \\
0 & & & 0 & & & 0\ 0\ 0 \\
 & & & & & & 0\ 0\ 1
\end{array}\right)$$

Её вполне положительность вытекает из критерия (3.4). Если $\Phi = diag\langle x, y, 1-x-y \rangle$, то $\Phi Ш = diag\langle \theta x, \theta y, 1-\theta x-\theta y \rangle$. Аналогично, $d\Phi = \begin{pmatrix} 0 & \gamma \\ \gamma & 0 \end{pmatrix} \oplus 0 \Rightarrow d\Phi Ш = \begin{pmatrix} 0 & \theta\gamma \\ \theta\gamma & 0 \end{pmatrix} \oplus 0 = \theta d\Phi$. Отсюда по предположенной монотонности $Q_{\Phi Ш}(d\Phi Ш) \leq Q_{\Phi}(d\Phi)$ следует (8.1), что и требовалось.

Ввиду однородности (7.10) функцию $c(x,y)$ естественно оценить на границе области задания, т.е. при $x+y=1$, $y = 1-x$, $x, y > 0$.

<u>Теорема 8.2.</u> Если риманова метрика ρ монотонна в категории, то выполнены неравенства

$$c(\tfrac{1}{2}, \tfrac{1}{2}) \leq c(x, 1-x) \leq c(\tfrac{1}{2}, \tfrac{1}{2}) \,/\, 4x(1-x) , \tag{8.3}$$

$$x > x' \geq \tfrac{1}{2} \implies c(x, 1-x) \geq c(x', 1-x') , \tag{8.4}$$

$$\frac{c(x, 1-x)}{1-x} \geq \frac{dc(x, 1-x)}{dx} \tag{8.5}$$

Доказательство всех этих неравенств проводится по одинаковому плану, подбором соответствующего марковского эндоморфизма совокупности $\mathfrak{S}(\mathfrak{L}(\mathcal{H}_2))$. Пусть $x > 1/2 > 1-x$. Рассмотрим эндоморфизм Π с четыре-матрицей второго порядка

$$\left(\begin{array}{cc|cc} 1-\vartheta & 0 & 0 & 1-\vartheta \\ 0 & \vartheta & \vartheta & 0 \\ \hline 0 & \vartheta & \vartheta & 0 \\ 1-\vartheta & 0 & 0 & 1-\vartheta \end{array}\right)$$

Она вполне положительна при всех $\vartheta \in [0,1]$. На операторы с матрицей $\mathrm{diag}\langle x, 1-x\rangle$ действует по формуле $[\mathrm{diag}\langle x,1-x\rangle]\Pi = \mathrm{diag}\langle x-\vartheta\Delta, 1-x+\vartheta\Delta\rangle$, где $\Delta = x-(1-x) = 2x-1 > 0$. по предположению. А для вещественного дифференциала $d\Phi = \begin{pmatrix} 0 & \gamma \\ \gamma & 0 \end{pmatrix}$, $d\Phi\,\Pi = d\Phi$. Так как при изменении ϑ от 0 до 1, $x' = x-\vartheta\Delta$ пробегает все значения в интервале $[x, 1-x]$, то отсюда, по монотонности $Q_{\Phi\Pi}(d\Phi\Pi) \leqslant Q_{\Phi}(d\Phi)$, следует (8.4). В частности, при $\vartheta = 1/2$, получаем левое неравенство (8.3).

Рассмотрим теперь при $0 < x < 1$ эндоморфизм Ж с матрицей

$$\left(\begin{array}{cc|cc} x & 0 & 0 & \sqrt{x(1-x)} \\ 0 & 1-x & \sqrt{x(1-x)} & 0 \\ \hline 0 & \sqrt{x(1-x)} & x & 0 \\ \sqrt{x(1-x)} & 0 & 0 & 1-x \end{array}\right)$$

и его действие на операторы и дифференциалы указанного выше вида. Имеем $\mathrm{diag}\langle 1/2, 1/2\rangle Ж = \mathrm{diag}\langle x, 1-x\rangle$, $d\Phi\, Ж = 2\sqrt{x(1-x)}\; d\Phi$. Выписанная четыре-матрица вполне положительна по критерию (3.4).

Отсюда, по монотонности метрики и свойству (7.10) получаем левое неравенство (8.3).

Наконец, рассмотрим эндоморфизм Щ с вполне положительной при $0 < \vartheta < 1$ матрицей

$$\left(\begin{array}{cc|cc} 1 & 0 & 0 & \sqrt{1-\vartheta} \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \vartheta & 0 \\ \sqrt{1-\vartheta} & 0 & 0 & 1-\vartheta \end{array}\right)$$

На интересующие нас операторы и дифференциалы он действует по формулам

$$[\mathrm{diag}\langle x, 1-x\rangle]\,Щ = \mathrm{diag}\langle x+\vartheta(1-x), (1-x)(1-\vartheta)\rangle,$$

$$d\Phi\,Щ = \sqrt{1-\vartheta}\; d\Phi$$

Отсюда по монотонности получаем

$$c(x, 1-x) \geqslant (1-\vartheta)\cdot c\bigl(x+\vartheta(1-x),\ 1-x-\vartheta(1-x)\bigr),$$

$$\vartheta\, c(x+\delta, 1-x-\delta) \geqslant c(x+\delta, 1-x-\delta) - c(x, 1-x),$$

где $\delta = \vartheta(1-x) > 0$. Поделив обе части на δ, приходим к неравенству

$$\frac{c(x+\delta,\ 1-x-\delta)}{1-x} \geqslant \frac{c(x+\delta, 1-x-\delta) - c(x, 1-x)}{\delta}.$$

Выражение справа неотрицательно. Выражение слева ограничено по (8.3) при любом $\delta < 1-x$, т.е. $\vartheta < 1$, и монотонно убывает по (8.4) при $\delta \downarrow 0$. Если предел справа существует, то он является правой производной и для нее выполнено неравенство (8.5). Так как функция $c(x, 1-x)$ монотонна, то она имеет производную почти всюду. В исключительных точках выполняется аналогичное неравенство для локальной константы Липшица. Теорема доказана.

Так как совокупность $\mathfrak{S}(\mathcal{L})$ выпукла, то можно рассматривать значения квадратичных форм $Q_\Phi(\cdot)$ не только от дифференциалов, но и от конечных приращений. Для классической фишеровской информации (т.е. для $\Delta\Phi$, коммутирующих с Φ)

$$Q_{\Phi\otimes\Psi}(\Delta[\Phi\otimes\Psi]) = Q_\Phi(\Delta\Phi) + Q_\Psi(\Delta\Psi) + \\ + Q_\Phi(\Delta\Phi)\cdot Q_\Psi(\Delta\Psi) . \tag{8.6}$$

<u>Теорема 8.3.</u> Если задающая монотонную риманову метрику квадратичная форма Q удовлетворяет обобщенному принципу сложения информации (8.6) при любых самосопряженных приращениях $\Delta\Phi$ и $\Delta\Psi$, то отвечающая ей функция $c(x,y)$ удовлетворяет функциональному уравнению

$$c(pu,qv) + c(pv,qu) = 2c(p,q)\cdot c(u,v) \tag{8.7}$$

при всех допустимых значениях аргументов

$$0 < p,q < p+q \leq 1 \quad , \quad 0 < u,v < u+v \leq 1 .$$

Доказательство. Пусть $\Phi = diag\langle p_1, p_2, \dots\rangle$, $\Psi = diag\langle q_1, q_2, \dots\rangle$, приращение $\Delta\Phi$ сводится к недиагональному возмущению $\Delta\varphi^i_j = \overline{\Delta\varphi^j_i}$, аналогично $\Delta\Psi$ $\Delta\psi^k_\ell = \overline{\Delta\psi^\ell_k}$. Поскольку

$$\Delta(\Phi\otimes\Psi) = \Phi\otimes(\Delta\Psi) + (\Delta\Phi)\otimes\Psi + (\Delta\Phi)\otimes(\Delta\Psi),$$

$$Q_{\Phi\otimes\Psi}(\Delta[\Phi\otimes\Psi]) = 2\sum_\alpha c(p_\alpha q_k, p_\alpha q_\ell)\left|p_\alpha \Delta\psi^k_\ell\right|^2 + \\ + 2\sum_\beta c(p_i q_\beta, p_j q_\beta)\left|q_\beta \Delta\varphi^i_j\right|^2 + 2\left[c(p_i q_k, p_j q_\ell) + \\ + c(p_i q_\ell, p_j q_k)\right]\cdot\left|\Delta\varphi^i_j\right|^2\cdot\left|\Delta\psi^k_\ell\right|^2 .$$

Так как $\sum p_\alpha = \sum q_\beta = 1$, то, вынося p_α и q_β по правилу (7.10) мы приводим первые два правых слагаемых к виду

$$2c(q_k, q_e)\,|\Delta\psi^k_e|^2 = Q_\Psi(\Delta\Psi)\,; \quad 2c(p_i, p_j)\,|\Delta\varphi^i_j|^2 = Q_\Phi(\Delta\Phi)\,.$$

Отсюда, если (8.6) выполняется, последняя сумма в квадратных скобках должна приводиться к произведению, т.е. соотношение (8.7) должно выполняться. Нетрудно выкладкой проверить, что условие (8.7) является и достаточным для выполнения (8.6).

<u>Лемма 8.4.</u> Функциональному уравнению (8.7) удовлетворяют функции

$$c(u, v) = c\cdot\left[u^\alpha v^{-\alpha} + u^{-\alpha} v^\alpha\right](uv)^\beta \tag{8.8}$$

при любых вещественных α и β . Они вещественно однородны в минус первой степени при $\beta = -1/2$. При $0 \leq \alpha \leq 1/2$ эти последние удовлетворяют неравенству (8.3)

Доказательство следует из тождества:

$$\left[u^\alpha v^{-\alpha} + u^{-\alpha} v^\alpha\right]\left[p^\alpha q^{-\alpha} + p^{-\alpha} q^\alpha\right] =$$

$$= \left[(pu)^\alpha (qv)^{-\alpha} + (pu)^{-\alpha}(qv)^\alpha\right] + \left[(pv)^\alpha (qu)^{-\alpha} + (pv)^{-\alpha}(qu)^\alpha\right]$$

В частных случаях имеем, см. [22] ,

$$Q_\Phi(d\Phi) = c\,tr\left[\Phi^{-1}(d\Phi)^2\right], \qquad \alpha = 1/2\,;$$

$$Q_\Phi(d\Phi) = c\,tr\left[(d\Phi)\,\Phi^{-1/2}\,(d\Phi)\,\Phi^{-1/2}\right], \quad \alpha = 0\,.$$

Однако, доказывать, что эти формы действительно порождают монотонные римановы метрики, мы не умеем. Все наши построения были необходимыми, но не достаточными. Однако, то, что их удается записать в вышеуказанных операторных терминах, доказывает их унитарную инвариантность.

9. Теоремы 5.1 и 6.2 показывают, что топология, порождаемая расстоянием по вариации, является самой слабой из метрических топологий, инвариантных относительно категории марковских отображений. В этом смысле они далеко обобщают результаты работ [23] и [3] . Воспользуемся ими сейчас для расширения нашего вывода [24] о некорректности классической обратной задачи теории вероятностей – задачи оценки распределения вероятностей случайной величины по последовательности ее независимых наблюдений $\xi_1, \dots, \xi_N, \dots$ при отсутствии всякой дополнительной априорной информации о этом распределении. Как мы показали [24] , если погрешность решения измерять вариацией разности $|P_N^* - P|$, где P_N^* есть оценка распределения P , то для любой последовательности решающих правил $\Pi_N(\xi_1, \dots, \xi_N)$, $N = 1, 2, \dots$, найдется такое распределение P , на котором последовательность Π_N не состоятельна. В силу неравенств (5.1) и (6.4) эта несостоятельность сохранится, будем ли мы измерять погрешность другой монотонной метрикой или же марковски монотонной функцией потерь.

Московский государственный университет им.М.В.Ломоносова
Институт прикладной математики им.М.В.Келдыша АН СССР

Литература

I. Ченцов Н.Н. Матем.заметки, I968, т.4, с.323-332.

2. Ченцов Н.Н. Статистические решающие правила и оптимальные выводы. М., Наука, I972.

3. Csiszar I., Studia Sci. Math. Hungar., 1967, v.2, p.329-339.

4. Ченцов Н.Н. ДАН СССР, I965, т. I64, с. 5II-5I4.

5. Morse N., Sacksteder R., Ann. Math. Statist., 1966, v.37 p.203-214.

6. Wald A., Ann. Math. Statist., 1939, v.10, p.299-326.

7. Cencov N.N., Math. Operations forsch. Statist., ser Statistics, 1978, v.9, p.267-276.

8. Neumann J. von. Grundlagen der Quantenmechanik, Berlin, 1932.

9. Emch G.G., Mathematical Methods in Statistical Mechanics and Quantum Field Theory, Wiley, 1972.

I0. Eilenberg S., Mac Lane S. Trans. Amer. Math. Soc., 1945, v.58, p.231-294.

II. Stinespring W.F., Proc. Amer. Math. Soc., 1955, v.6, p.211-216.

I2. Наймарк М.А. Изв.АН СССР, сер. матем., I940, т.4, с.277-3I8.

I3. Холево А.С. Вероятностные и статистические аспекты квантовой теории, М., Наука, I980.

I4. Effros E.G., Stormer E., Math. Scand., 1979, v.45, p.127-138.

I5. Cencov N.N., 12-th Europ. Meeting of Statisticians, Varna, 1979.

I6. Морозова Е.А., Ченцов Н.Н. Препринт Ин. прикл.матем. им. М.В.Келдыша АН СССР, I98I, № I.

I7. Cencov N.N., IV USSR-Japan Symposium on Probability Theory and Mathematical Statistics, Abstracts, 1982, v.1, p.162-164.

I8. Morozova E.A., Cencov N.N., Lecture Notes in Mathematics, N 550, Springer, 1976, p.379-418.

I9. Сарымсаков Т.А. Введение в квантовую теорию вероятностей, Ташкент, ФАН, 1985.

20. Kraus K., Ann. Phys., 1971, v.64, p.331-335.

2I. Morozova E.A., Cencov N.N., Fourth International Vilnius Conference on Probability Theory and Mathematical Statistics, Abstracts, 1985, v.2, p.218-220.

22. Морозова Е.А., Ченцов Н.Н. УМН, 1985, т.40, с. I89-I90.

23. Ченцов Н.Н. УМН, I967, т.22, с. I78-I80.

24. Ченцов Н.Н. Теория вероятн. примен., I98I, т.26, с.I5-3I.