PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 3, Fasc. 1 (1982), p. 113-126

TESTING STATISTICAL HYPOTHESES IN QUANTUM THEORY

BY

A. S. HOLEVO (Moscow)

Abstract. The paper presents an account of quantum hypotheses testing theory and the new contributions clarifying the role of orthogonal resolutions of identity (simple measurements) in the quantum Bayes problem. It is shown that the maximum likelihood measurement is simple for the family of states with linearly independent ranges. This is an extension of Kennedy's result for the pure states. A counterexample to an old physical conjecture is constructed showing that in the Bayes problem with the total number of decisions less than or equal to the dimension of the underlying Hilbert space the Bayes measurement may not be simple. However, it is shown that in the infinite-dimensional case there always exists an ε-Bayes simple measurement.

О ПРОВЕРКЕ СТАТИСТИЧЕСКИХ ГИПОТЕЗ В КВАНТОВОЙ ТЕОРИИ

А. С. ХОЛЕВО (Москва)

1. В этой работе рассматривается квантовая задача различения гипотез, математическая формулировка которой состоит в следующем (мотивировку и более подробное обсуждение см. в [3], [5], [7]). В сепарабельном гильбертовом пространстве $\mathscr{H}$ заданы операторы плотности (о. п.) S_j $(j = 1, \ldots, m)$, представляющие конкурирующие гипотезы о с о с т о я н и и рассматриваемой квантовой системы. *Оператор плотности* S — это положительный оператор с единичным следом:

$$S \geqslant 0, \quad \mathrm{Tr} S = 1. \tag{1}$$

Решение о выборе той или иной гипотезы принимается на основании *квантового измерения*, которое описывается разложением единицы в $\mathscr{H}$, т. е. набором $\boldsymbol{M}$ эрмитовых операторов M_k, $k=1,\dots,m$, удовлетворяющим условиям

$$(2) \qquad M_k \geqslant 0, \quad k=1,\dots,m, \sum_{k=1}^{m} M_k = I,$$

где I — единичный оператор в $\mathscr{H}$. При этом вероятность принять решение k, если состояние описывается о. п. S, считается равной

$$\mu_S^{\boldsymbol{M}}(k) = \mathrm{Tr} S M_k.$$

В силу (2) величины $\mu_S^{\boldsymbol{M}}(k)$, $k=1,\dots,m$, образуют распределение вероятностей. Таким образом, вероятность принять решение k, если истинным состоянием является S_j, равна

$$(3) \qquad p_j(k) = \mathrm{Tr}\, S_j M_k.$$

Если $j \neq k$, то принятое решение является ошибочным, и задача заключается в том, чтобы выбрать измерение $\boldsymbol{M}$ в каком-то смысле наилучшим образом.

Как и в классической статистике, обычно формируется тот или иной функционал от вероятностей ошибок (3) и ищется оптимальное измерение, дающее ему подходящий экстремум.

При байесовском подходе, которым мы и ограничимся, задаются априорные вероятности гипотез π_j, $j=1,\dots,m$, и функция потерь $W_j(k)$, $j,k=1,\dots,m$. Байесовский риск определяется обычной формулой

$$(4) \qquad \mathscr{R}\{\boldsymbol{M}\} = \sum_{j=1}^{m} \pi_j \sum_{k=1}^{m} W_j(k) p_j(k),$$

в которой, однако, вероятности $p_j(k)$ зависят от $\boldsymbol{M}$ по квантово-механической формуле (3). Измерение, минимизирующее $\mathscr{R}\{\boldsymbol{M}\}$, называется *байесовским*. Часто рассматривается случай *простой функции потерь* $W_j(k) = 1-\delta_{jk}$. Минимизация соответствующего байесовского риска эквивалентна максимизации средней вероятности правильного решения

$$\mathscr{P}\{\boldsymbol{M}\} = \sum_{j=1}^{m} \pi_j p_j(j).$$

Если $\pi_j \equiv 1/m$, то

$$(5) \qquad \mathscr{P}\{\boldsymbol{M}\} = \frac{1}{m} \sum_{j=1}^{m} p_j(j).$$

Измерение, максимизирующее (5), мы называем *измерением максимального правдоподобия* (объяснение такой терминологии можно найти в [7]).

Сформулируем некоторые общие результаты о байесовской задаче, которые понадобятся в дальнейшем.

Обозначим $\mathfrak{M}$ выпуклое множество всех измерений. Последовательность $\{M^{(n)}\} \subset \mathfrak{M}$ сходится к M, если $\mu_S^{M^{(n)}}(k) \to \mu_S^M(k)$, $k = 1, \ldots, m$, для любого о. п. S. Используя теорему Банаха-Алаоглу (см., например, [2]), можно показать, что $\mathfrak{M}$ является компактным топологическим пространством. Очевидно, что байесовский риск (4) является непрерывным аффинным функционалом на $\mathfrak{M}$. Поэтому *байесовское измерение всегда существует, и может быть найдено среди крайних точек множества* $\mathfrak{M}$ [5].

Используя (4) и (3), получаем

$$\mathscr{R}\{M\} = \mathrm{Tr} \sum_{k=1}^{m} \hat{W}_k M_k, \tag{6}$$

где $\hat{W}_k = \sum_{j=1}^{m} \pi_j S_j W_j(k)$ — операторная „апостериорная функция потерь". *Для того, чтобы измерение* $M^{(0)} = \{M_k^{(0)}\}$ *было оптимальным, необходимо и достаточно, чтобы оператор*

$$\Lambda = \sum_{k=1}^{m} \hat{W}_k M_k^{(0)} \tag{7}$$

был эрмитов и удовлетворял неравенствам

$$\Lambda \leqslant \hat{W}_k, \quad k = 1, \ldots, m. \tag{8}$$

При этом выполняется

$$(\Lambda - \hat{W}_k) M_k^{(0)} = 0, \quad k = 1, \ldots, m \tag{9}$$

(см. [5], [6], [8]). Для случая (5) имеет место представление

$$\mathscr{P}\{M\} = \frac{1}{m} \sum_{j=1}^{m} S_j M_j, \tag{6'}$$

а соотношения (7)-(9) сводятся к

$$\Lambda = \sum_{j=1}^{m} S_j M_j^{(0)}, \tag{7'}$$

$$\Lambda \geqslant S_j, \quad j = 1, \ldots, m, \tag{8'}$$

$$(\Lambda - S_j) M_j^{(0)} = 0, \quad j = 1, \ldots, m. \tag{9'}$$

Оптимальное измерение, вообще говоря, неединственно; однако оператор Λ, вычисляемый по формуле (7) или (7'), одинаков для всех оптимальных измерений [6].

Измерение называется *простым*, если оно описывается ортогональным разложением единицы $\boldsymbol{E} = \{E_k\}$:

$$(10) \qquad E_j E_k = \delta_{jk} E_k.$$

Это означает, что E_k, $k = 1, \ldots, m$, являются ортогональными проекторами на взаимно-перпендикулярные подпространства в $\mathscr{H}$. Теорема Наймарка [1] утверждает, что *всякое разложение единицы $\{M_k\}$ в гильбертовом пространстве $\mathscr{H}$ может быть продолжено до ортогонального разложения единицы $\{E_k\}$ в гильбертовом пространстве $\tilde{\mathscr{H}} \supset \mathscr{H}$*, так что

$$M_k = PE_kP, \quad k = 1, \ldots, m,$$

где P проектор из $\tilde{\mathscr{H}}$ на $\mathscr{H}$. Таким образом, измерение, задаваемое неортогональным разложением единицы, можно рассматривать как простое измерение над некоторым расширением исходной системы [5]. В этом смысле можно провести аналогию между простыми измерениями и детерминированными стратегиями и, соответственно, между измерениями и рандомизированными стратегиями в классической статистике. Следующий пример, однако, показывает, что в квантовой байесовской задаче нельзя ограничиться рассмотрением простых измерений (тогда как в классической задаче всегда существует детерминированная байесовская стратегия).

Пусть $\mathscr{H}$ двумерное гильбертово пространство с фиксированным ортонормированным базисом, и о. п. $\{S_j\}$ задаются в этом базисе матрицами

$$(11) \qquad S_j = \frac{1}{2}\begin{bmatrix} 1 & e^{-2\pi ij/m} \\ e^{2\pi ij/m} & 1 \end{bmatrix}, \quad j = 1, \ldots, m.$$

Состояния S_j описывают „частицу со спином $1/2$, поляризованную в направлениях, лежащих в одной плоскости и образующих правильный m-угольник". В силу того, что $\sum_{j=1}^{m} e^{-2\pi ij/m} = 0$, имеем

$$\sum_{j=1}^{m} S_j = \frac{m}{2} I.$$

Поэтому операторы

$$(12) \qquad M_j = \frac{2}{m} S_j, \quad j = 1, \ldots, m,$$

образуют разложение единицы в $\mathscr{H}$, которое ортогонально лишь при $m = 2$. Это разложение единицы является измерением максимального правдоподобия для семейства (11), т. к. оператор

$$\Lambda = \sum_j S_j M_j = \frac{2}{m} \sum_j S_j^2 = \frac{2}{m} \sum_j S_j = I$$

эрмитов и удовлетворяет условию (8). Как мы увидим в п. 3, существуют и другие измерения максимального правдоподобия, однако при m нечетном все они описываются неортогональными разложениями единицы. Так как $\mathscr{H}$ конечномерно, то множество ортогональных разложений единицы $\mathfrak{E}$, определяемое конечным набором равенств (10), является замкнутым подмножеством компактного множества $\mathfrak{M}$. Следовательно

$$\max_{\mathfrak{M}} \mathscr{P}\{M\} > \max_{\mathfrak{E}} \mathscr{P}\{M\} \quad \text{при } m = 2k+1,\ k \geqslant 1.$$

В свете изложенного представляет интерес отыскание условий, при которых оптимальное измерение является простым; эта дополнительная информация значительно сужает класс измерений, в котором ведется поиск экстремума. За исключением очевидного результата для коммутирующих о. п., который сводится к классической байесовской задаче (см., например, [3]), в этом направлении имелось единственное продвижение — теорема Кеннеди, относящаяся к случаю „чистых линейно независимых" состояний. П. 2 посвящен распространению этого результата на более общие о. п. Важную роль здесь играет новое понятие приводящего подпространства. В физической литературе высказывалось предположение (мотивируемое отчасти теоремой Кеннеди), что при числе гипотез $m \geqslant \dim \mathscr{H}$ байесовское измерение всегда может быть выбрано простым (см. например, [4]). В п. 3 мы строим опровергающий пример, который является видоизменением конструкции (11). Однако, оказывается справедливым более слабое утверждение: если $\dim \mathscr{H} = \infty$, то всегда

$$\inf_{\mathfrak{E}} \mathscr{R}\{M\} = \min_{\mathfrak{M}} \mathscr{R}\{M\}.$$

Это связано с тем, что в случае $\dim \mathscr{H} = \infty$ простые измерения образуют п л о т н о е п о д м н о ж е с т в о в множестве всех измерений.

2. Пусть ψ_j, $j = 1, \ldots, m$, линейно независимые единичные векторы в $\mathscr{H}$, порождающие подпространство $\mathscr{H}_0$ и S_j, $j = 1, \ldots, m$, соответствующие чистые состояния, т. е. $S_j = E[\psi_j]$, где $E[\psi]$ — оператор ортогонального проектирования на вектор ψ. Не ограничивая общности, задачу различения состояний $S_j, j = 1, \ldots, m$, можно рассматривать в пространстве $\mathscr{H}_0$ (точный смысл этого утверждения будет выяснен далее). Теорема Кеннеди (см. [3], п. IV,1c) утверждает, что при простой функции потерь байесовское измерение (в частности измерение максимального правдоподобия) для семейства $\{S_j\}$ в $\mathscr{H}_0$ является простым и имеет вид $M_j = E[\varphi_j]$, $j = 1, \ldots, m$, где $\{\varphi_j\}$ — некоторый ортонормированный базис в $\mathscr{H}_0$.

Прежде чем сформулировать обобщение этого результата, введем понятие приводящего подпространства. Рассмотрим функционал вида (6). (Замкнутое) подпространство $\mathscr{H}'$ назовем *приводящим*, если для любого измерения $\{M_k\}$

$$\text{(13)}\qquad \operatorname{Tr}\sum_{k=1}^{m}\hat{W}_kP'M_kP' = \operatorname{Tr}\sum_{k=1}^{m}\hat{W}_kM_k+c,$$

где P' — проектор на $\mathscr{H}'$, а $c=c(\mathscr{H}')$. Если $\boldsymbol{M}^{(o)}=\{M_k^{(o)}\}$ — оптимальное измерение для функционала (6), а $\mathscr{H}'$ — приводящее подпространство, то $\boldsymbol{M}^{(o)\prime}=\{P'M_k^{(o)}P'\}$ — оптимальное измерение для функционала $\operatorname{Tr}\sum_k P'\hat{W}_kP'M_k$ в пространстве $\mathscr{H}'$. Обратно, если $\boldsymbol{M}^{(o)\prime}$ — оптимальное измерение в пространстве $\mathscr{H}'$, то любое продолжение $\boldsymbol{M}^{(o)}$ измерения $\boldsymbol{M}^{(o)\prime}$ является оптимальным измерением для исходного функционала в $\mathscr{H}$. (Отметим, что продолжение всегда существует; можно положить, например, $M_k^{(o)}=M_k^{(o)\prime}\oplus N_k$, где $\{N_k\}$ — произвольное разложение единицы в $\mathscr{H}\ominus\mathscr{H}'$.) В этом смысле задача оптимизации на пространстве $\mathscr{H}$ сводится к соответствующей задаче на приводящем подпространстве $\mathscr{H}'$.

Предложение 1. *Существует минимальное приводящее подпространство $\mathscr{H}_0$ (т. е. содержащееся в любом другом приводящем подпространстве). Это есть замыкание множества векторов ψ, которое задается одним из следующих эквивалентных соотношений:*

$$\text{(14)}\qquad \psi=\sum_k\hat{W}_k\psi_k,\quad \text{где } \psi_k\in\mathscr{H} \text{ и } \sum_k\psi_k=0;$$

$$\text{(14')}\qquad \psi=\sum_k(\hat{W}_k-\hat{W}_j)\psi_k,\quad \text{где } j \text{ — фиксировано и } \psi_k\in\mathscr{H};$$

$$\text{(14'')}\qquad \psi=\sum_{k,j}(\hat{W}_k-\hat{W}_j)\psi_{kj},\quad \text{где } \psi_{kj}\in\mathscr{H}.$$

Доказательство. Обозначим $\mathscr{L},\mathscr{L}_j',\mathscr{L}''$ множества векторов, задаваемые соответственно, соотношениями (14), (14'), (14''). Очевидно, что $\mathscr{L}\subset\mathscr{L}_j'\subset\mathscr{L}''$. Пусть $\psi\in\mathscr{L}''$; тогда

$$\psi=\sum_k\hat{W}_k\sum_j(\psi_{kj}-\psi_{jk}).$$

Полагая $\psi_k=\sum_j(\psi_{kj}-\psi_{jk})$, имеем

$$\sum_k\psi_k=0,$$

так что $\psi\in\mathscr{L}$. Итак, $\mathscr{L}''\subset\mathscr{L}$ и $\mathscr{L}=\mathscr{L}_j'=\mathscr{L}''$. Обозначим $\mathscr{H}_0$ подпространство $\mathscr{H}$, которое является замыканием этого множества. Из (14'') следует, что

$$\text{(15)}\qquad \mathscr{H}_0=\bigvee_{j,k}\mathscr{R}(\hat{W}_j-\hat{W}_k),$$

где $\mathscr{R}(A)$ — область значений оператора A, а символ $\bigvee$ обозначает замкнутую линейную оболочку. Поэтому проектор P_0 на $\mathscr{H}_0$ является минимальным проектором P, удовлетворяющим соотношению

$$P(\hat{W}_i-\hat{W}_j) = \hat{W}_i-\hat{W}_j, \quad i,j=1,\ldots,m. \tag{16}$$

Пусть $\mathscr{H}'$ — произвольное приводящее подпространство. Мы покажем, что проектор $P=P'$ удовлетворяет соотношению (16) и тем докажем предложение. Из (13) следует

$$\operatorname{Tr}\sum_k \hat{W}_k(M_k-P'M_kP') = c.$$

Фиксируем j и заметим, что $M_j = I-\sum_{k:\,k\neq j} M_k$; следовательно

$$\operatorname{Tr}\sum_{k:\,k\neq j}(\hat{W}_k-\hat{W}_j)(M_k-P'M_kP') = c'$$

для любого измерения $\{M_k\}$. Полагая $M_k=0$ для $k\neq j$, так что $M_j=I$, получаем $c'=0$. Полагая $M_k=0$ для $k\neq i,j$, получаем

$$0=\operatorname{Tr}(\hat{W}_i-\hat{W}_j)(M_i-P'M_iP') = \operatorname{Tr}[(\hat{W}_i-\hat{W}_j)-P'(\hat{W}_i-\hat{W}_j)P']M_i$$

для любого M_i $(0\leqslant M_i\leqslant I)$. Отсюда

$$\hat{W}_i-\hat{W}_j = P'(\hat{W}_i-\hat{W}_j)P'. \tag{17}$$

Умножая это на P' и сравнивая результирующее равенство с исходным, получаем $\hat{W}_i-\hat{W}_j = P'(\hat{W}_i-\hat{W}_j)$. Предложение доказано.

Учитывая, что $\mathscr{R}(A) = \mathscr{W}(A)^{\perp}$, где $\mathscr{W}(A)$ — нулевое подпространство оператора A, а также соотношение двойственности в решетке подпространств получаем эквивалентную форму соотношения (15):

$$\mathscr{H}_0 = [\bigwedge_{j,k}\mathscr{W}(\hat{W}_j-\hat{W}_k)]^{\perp},$$

где $\bigwedge$ обозначает пересечение подпространств. Очевидно, что

$$\bigwedge_{j,k}\mathscr{W}(\hat{W}_j-\hat{W}_k) = \{\psi\colon \hat{W}_1\psi=\ldots=\hat{W}_m\psi\}.$$

Обозначая последнее подпространство $[\hat{W}_1=\ldots=\hat{W}_m]$, имеем

$$\mathscr{H}_0 = [\hat{W}_1=\ldots=\hat{W}_m]^{\perp}. \tag{18}$$

Из (17) следует, что для любого приводящего подпространства $\mathscr{H}'$ оператор $\hat{W}_j-P'\hat{W}_jP'$ одинаков для всех $j=1,\ldots,m$; обозначая его $\hat{W}'$, имеем $\hat{W}_j = P'\hat{W}_jP'+\hat{W}'$, так что в формуле (13) $c(\mathscr{H}') = \operatorname{Tr}\hat{W}'$.

В условиях теоремы Кеннеди минимальное приводящее подпространство для $\mathscr{P}\{M\} = \sum \pi_j p_j(j)$ совпадает с линейной оболочкой векторов ψ_j, $j = 1, \ldots, m$. Это следует из формулы (17) и того факта, что $\mathscr{R}(S_j - S_k) = \mathscr{R}(E[\psi_j] - E[\psi_k])$ есть двумерное подпространство, натянутое на ψ_j, ψ_k.

Следующее предложение обобщает результат Кеннеди на „смешанные" (т. е. не обязательно чистые) состояния $\{S_j\}$:

Предложение 2. *Пусть области значений $\mathscr{R}(S_j)$, $j = 1, \ldots, m$, линейно независимы. Тогда, при простой функции потерь, байесовское измерение для семейства о. п. $\{S_j\}$ единственно и является простым на минимальном приводящем подпространстве $\mathscr{H}_0$.*

Линейная независимость означает, что из $\sum_j \varphi_j = 0$, $\varphi_j \in \mathscr{R}(S_j)$, $j = 1, \ldots, m$, следует $\varphi_j = 0$, $j = 1, \ldots, m$, или что из $\sum_j S_j\psi_j = 0$ $(\psi_j \in \mathscr{H})$ следует $S_j\psi_j = 0$, $= 1, \ldots, m$. В силу линейной независимости,

$$\mathscr{R}(S_j - S_k) = \mathscr{R}(S_j) \vee \mathscr{R}(S_k),$$

так что

$$\mathscr{H}_0 = \bigvee_j \mathscr{R}(S_j).$$

Теорема Кеннеди является частным случаем предложения 2 при $\dim \mathscr{R}(S_j) \equiv 1$. Одно из его применений будет приведено в п. 3.

Мы получим предложение 2 как следствие более абстрактной теоремы, относящейся к общей байесовской задаче. Заметим, что оптимальное измерение обладает инвариантностью по отношению к преобразованию $\hat{W}_k \to \hat{W}_k + \hat{W}$, $k = 1, \ldots, m$, так как

$$\mathrm{Tr} \sum_k (\hat{W}_k + \hat{W}) M_k = \mathrm{Tr} \sum_k \hat{W}_k M_k + \mathrm{Tr}\,\hat{W}. \tag{19}$$

Эта инвариантность учитывается в формулировке следующей теоремы (тогда как условия предложения 2 ею не обладают):

Теорема 1. *Пусть $\hat{W}_k$, $k = 1, \ldots, m$, таковы, что существует эрмитов оператор $\tilde{\Lambda}$ со следующими свойствами:*

(а) $\tilde{\Lambda} \geqslant \hat{W}_k$, $k = 1, \ldots, m$;

(б) *из* $\sum_k \psi_k = 0$, $\sum_k \hat{W}_k\psi_k = 0$ $(\psi_k \in \mathscr{H})$ *следует* $\hat{W}_k\psi_k = \tilde{\Lambda}\psi_k$, $k = 1, \ldots, m$.

Тогда байесовское измерение единственно и является простым на минимальном приводящем подпространстве $\mathscr{H}_0$.

Предложение 2 получается, если положить в этой теореме $\hat{W}_k = -\pi_k S_k$, $\tilde{\Lambda} = 0$ и пренебречь условием $\sum_k \psi_k = 0$.

Доказательство. Покажем сначала, что оператор $\tilde{\Lambda}-\Lambda$ не вырожден на $\mathscr{H}_0$. В силу (8) и (а), $\tilde{\Lambda}-\Lambda \geqslant \tilde{\Lambda}-\hat{W}_k \geqslant 0$, $k = 1, \ldots, m$. Поэтому

$$\mathscr{W}(\tilde{\Lambda}-\Lambda) \subseteq \bigwedge_k \mathscr{W}(\tilde{\Lambda}-\hat{W}_k) \subseteq [\hat{W}_1 = \ldots = \hat{W}_m]$$

и, согласно (18), $\mathscr{W}(\tilde{\Lambda}-\Lambda) \wedge \mathscr{H}_0 = 0$.

Будем считать далее, что $\mathscr{H} = \mathscr{H}_0$; тогда $\mathscr{W}(\tilde{\Lambda}-\Lambda) = 0$. Докажем единственность. Пусть $\{M_k^{(1)}\}$, $\{M_k^{(2)}\}$ два байесовских измерения; поскольку, как отмечалось в п. 1, оператор Λ для них одинаков, то полагая $N_k = M_k^{(1)}-M_k^{(2)}$, получаем из (9)

$$(\Lambda-\hat{W}_k)N_k = 0, \quad k = 1, \ldots, m. \tag{20}$$

Суммируя по k и учитывая, что $\sum_k N_k = 0$, получаем

$$\sum_k \hat{W}_k N_k = 0.$$

Полагая $\psi_k = N_k\psi$, $\psi \in \mathscr{H}$, имеем

$$\sum_k \psi_k = 0, \qquad \sum_k \hat{W}_k \psi_k = 0,$$

откуда, согласно (б), $\hat{W}_k\psi_k = \tilde{\Lambda}\psi_k$, т. е. $\hat{W}_k N_k = \tilde{\Lambda}N_k$. Сравнивая с (20), получаем $(\tilde{\Lambda}-\Lambda)N_k = 0$, откуда, учитывая, что $\mathscr{W}(\tilde{\Lambda}-\Lambda) = 0$, следует $N_k = 0$, $k = 1, \ldots, m$.

Пусть $M^{(0)} = \{M_k^{(0)}\}$ байесовское измерение в $\mathscr{H}_0$. Из (7) и (9) получаем

$$\Big(\hat{W}_k - \sum_j \hat{W}_j M_j^{(0)}\Big)M_k^{(0)} = 0,$$

откуда

$$\sum_j \hat{W}_j(\delta_{jk}M_j^{(0)} - M_j^{(0)}M_k^{(0)}) = 0, \qquad k = 1, \ldots, m.$$

Пусть ψ — произвольный вектор из $\mathscr{H}_0$. Полагая

$$\psi_j = (\delta_{jk}M_j^{(0)} - M_j^{(0)}M_k^{(0)})\psi,$$

имеем $\sum_j \psi_j = 0$ и $\sum_j \hat{W}_j\psi_j = 0$, откуда, согласно (б), $\hat{W}_j\psi_j = \tilde{\Lambda}\psi_j$, т.е.

$$(\tilde{\Lambda}-\hat{W}_j)(\delta_{jk}M_j^{(0)} - M_j^{(0)}M_k^{(0)})\psi = 0, \qquad \psi \in \mathscr{H}_0.$$

Учитывая еще раз (9), получаем

$$(\tilde{\Lambda}-\Lambda)(\delta_{jk}M_j^{(0)} - M_j^{(0)}M_k^{(0)}) = 0, \qquad j,\ k = 1, \ldots, m,$$

откуда, в силу того что $\mathscr{W}(\tilde{\Lambda}-\Lambda)=0$, следует $\delta_{jk}M_j^{(0)}-M_j^{(0)}M_k^{(0)}=0$. Таким образом, $\{M_k^{(0)}\}$ — простое измерение, и теорема доказана.

3. Рассмотрим пространство $\mathscr{H}=\mathscr{H}^{(2)}\otimes\mathscr{H}'$, где $\mathscr{H}^{(2)}$ — двумерное гильбертово пространство, а $\mathscr{H}'$ имеет произвольную конечную или счетную размерность. Таким образом, $\dim\mathscr{H}=2\dim\mathscr{H}'$. Рассмотрим в $\mathscr{H}$ семейство о. п. вида

$$S_j=S_j^{(2)}\otimes S',\quad j=1,\dots,m, \tag{21}$$

где $S_j^{(2)}$ — о. п. в $\mathscr{H}^{(2)}$, которые даются формулой (11), а S' — произвольный невырожденный о. п. в $\mathscr{H}'$. Опишем все измерения максимального правдоподобия для семейства (21). Одно такое измерение получается очевидным образом из (12):

$$M_j=\frac{2}{m}S_j^{(2)}\otimes I',\quad j=1,\dots,m,$$

где I' единичный оператор в $\mathscr{H}'$. Отсюда получаем $\Lambda=I^{(2)}\otimes I'=I$. Поскольку оператор Λ для всех измерений максимального правдоподобия одинаков, то для их нахождения можно использовать уравнения (9′) с $\Lambda=I$. Из них легко следует, что

$$M_j=2S_j^{(2)}\otimes M_j',\quad j=1,\dots,m, \tag{22}$$

где M_j' — оператор в $\mathscr{H}'$. Найдем условия, которым должны удовлетворять операторы $\{M_j'\}$, чтобы $\{M_j\}$ было разложением единицы. Условие $M_j\geqslant 0$ равносильно тому что $M_j'\geqslant 0$. Используя матричное представление о. п. $S_j^{(2)}$, находим

$$\sum_j M_j=\begin{bmatrix}1&0\\0&1\end{bmatrix}\otimes\sum_j M_j'+\begin{bmatrix}0&0\\1&0\end{bmatrix}\otimes\sum_j e^{2\pi ij/m}M_j'+\begin{bmatrix}0&1\\0&0\end{bmatrix}\otimes\sum_j e^{-2\pi ij/m}M_j'.$$

Отсюда видно, что $\sum_j M_j=I$ тогда и только тогда, когда

$$\sum_{j=1}^m M_j'=I,\quad \sum_{j=1}^m e^{2\pi ij/m}M_j'=0. \tag{23}$$

Таким образом, *всякое измерение максимального правдоподобия для семейства* (21) *имеет вид* (22), *где* $\{M_j'\}$ — *произвольное разложение единицы в* $\mathscr{H}'$, *удовлетворяющее дополнительному условию* (23). Пусть $m=2k+1$, $k\geqslant 1$; тогда $S_i^{(2)}S_j^{(2)}\neq 0$ при $i\neq j$ и поэтому $M_iM_j=0$ тогда и только тогда, когда $M_i'M_j'=0$, т. е. когда $\{M_j'\}$ является ортогональным разложением единицы. Но это не совместимо с (23). Таким образом, при $m\geqslant 3$ любое измерение максимального правдоподобия задается неортогональным разложением единицы, тогда как размерность $\mathscr{H}$ может быть сколь угодно большой, и даже бесконечной.

В свете этого следующий результат кажется несколько неожиданным. Пусть $\varepsilon > 0$; назовем измерение $M^{(\varepsilon)}$ *ε-байесовским*, если

$$\mathscr{R}\{M^{(\varepsilon)}\} \leqslant \min_{\mathfrak{M}} \mathscr{R}\{M\}+\varepsilon.$$

Предложение 3. *Пусть* $\dim \mathscr{H} = \infty$. *Тогда для любого* $\varepsilon > 0$ *существует простое ε-байесовское измерение.*

Как будет показано ниже, при условии $\dim \mathscr{H} = \infty$ простые измерения образуют плотное (в смысле сходимости, определенной в п. 1) подмножество в множестве всех измерений. Предложение 3 следует из этого факта и непрерывности байесовского риска.

Мы используем предложения 2 и 3, чтобы обсудить еще одну догадку, высказанную в работе Хелстрома [4]. Объяснение терминов из теории оптической связи можно найти, например, в [3]. Рассматривается задача различения о. п.

$$S_j = W(\mu_j)S(0)W(\mu_j)^*, \quad j = 1, \ldots, m, \tag{24}$$

представляющих „когерентные сигналы на фоне теплового шума". Здесь $W(\mu)$ — унитарный „оператор сдвига", μ_j — комплексные амплитуды сигналов, а $S(0)$ — о. п. теплового шума:

$$S(0) = \sum_{n=0}^{\infty}(1-v)v^n E(n) \quad (0 < v < 1),$$

где $E(n)$ — проектор на n-й собственный вектор оператора числа квантов. Поскольку точное решение не известно, в [4] предлагается приближенный метод максимизации функционала $\mathscr{P}\{M\}$, использующий конечномерные аппроксимации о. п. (24). В пределах точности вычислений результат приближенной процедуры дается простым измерением. На основании этого факта, а также рассмотрения предельных случаев $v \to 0$, $v \to 1$, в [4] выдвигается предположение, что измерение максимального правдоподобия для семейства (24) является простым при любом v, $0 \leqslant v \leqslant 1$.

Рассмотрим операторы

$$S_j^{(N)} = W(\mu_j)\Big[\sum_{n=0}^{N}(1-v)v^n E(n)\Big]W(\mu_j)^*, \quad j = 1, \ldots, m,$$

которые также дают конечномерное приближение к семейству (24). Если N такое, что

$$\sum_{n=N+1}^{\infty}(1-v)v^n = v^N \leqslant \varepsilon,$$

то для любого измерения $\boldsymbol{M}$

$$0 \leqslant \mathscr{P}\{\boldsymbol{M}\} - \mathscr{P}^{(N)}\{\boldsymbol{M}\} \leqslant \varepsilon, \quad \text{где } \mathscr{P}^{(N)}\{\boldsymbol{M}\} = \frac{1}{m} \operatorname{Tr} \sum_j S_j^{(N)} M_j.$$

Отсюда также следует, что

$$0 \leqslant \max \mathscr{P}\{\boldsymbol{M}\} - \max \mathscr{P}^{(N)}\{\boldsymbol{M}\} \leqslant \varepsilon.$$

Таким образом, измерение, максимизирующее $\mathscr{P}^{(N)}\{\boldsymbol{M}\}$, является ε-оптимальным для исходной задачи. Используя предложение 2, мы можем легко показать, что это измерение — простое. Это следует из того, что любая конечная совокупность собственных векторов о. п. (24) является линейно независимой. В самом деле, в ,,представлении по когерентным состояниям" эти векторы задаются функциями

$$(n!)^{-1/2}(z-\mu_j)^n \exp\left(z\bar{\mu}_j - \tfrac{1}{2}|\mu_j|^2\right), \quad n = 0, 1, \ldots; \; j = 1, \ldots, m,$$

которые, очевидно, линейно независимы. К сожалению, предложение 2 не дает возможности сделать какой-либо вывод об измерении максимального правдоподобия для исходного семейства (24); состояния $\{S_j\}$ не удовлетворяют условиям этого предложения, так как $\mathscr{R}\{S_j\}$, $j = 1, \ldots, m$, имеют непустое (и даже плотное в $\mathscr{H}$) пересечение. Это следует из формулы Иошитани (соотношение (V.4.16) в [3]), которая показывает, что область значений любого оператора вида (24) содержит векторы всех ,,когерентных состояний". Однако предложение 3 показывает, что наличие ε-оптимальных простых измерений не может служить аргументом в пользу того, что оптимальное измерение является простым.

Желая получить обобщение предложения 3 на случай байесовской задачи с произвольным множеством решений (в частности, на байесовскую задачу оценивания), рассмотрим сепарабельное метрическое пространство U с σ-алгеброй борелевских множеств $\mathscr{B}(U)$. Разложение единицы на U в гильбертовом пространстве $\mathscr{H}$ может быть определено как семейство эрмитовых операторов $\boldsymbol{M} = \{M(B);\, B \in \mathscr{B}(U)\}$ такое, что для любого о. п. S в $\mathscr{H}$ функция множеств $\mu_S^{\boldsymbol{M}}(B) = \operatorname{Tr} SM(B)$, $B \in \mathscr{B}(U)$, является вероятностной мерой. Разложение единицы $\boldsymbol{E} = \{E(B)\}$ называется *ортогональным*, если $E(B_1)E(B_2) = E(B_1 \cap B_2)$, $B_1, B_2 \in \mathscr{B}(U)$. Обозначим $\mathfrak{M}(U)$ множество всех разложений единицы на U. Последовательность $\{\boldsymbol{M}^{(n)}\} \subset \mathfrak{M}(U)$ сходится к $\boldsymbol{M}$, если для любого о. п. S

$$\operatorname{var}\left(\mu_S^{\boldsymbol{M}^{(n)}} - \mu_S^{\boldsymbol{M}}\right) \to 0.$$

Если U конечное множество, то $\mathscr{B}(U)$ σ-алгебра всех конечных подмножеств U. Тогда

$$\mu(B) = \sum_{j \in B} M_j, \quad \operatorname{var}\left(\mu_S^{\boldsymbol{M}^{(n)}} - \mu_S^{\boldsymbol{M}}\right) = \sum_{j \in U} |\mu_S^{\boldsymbol{M}^{(n)}}(j) - \mu_S^{\boldsymbol{M}}(j)|,$$

и сходимость по вариации равносильна сходимости, введенной в п. 1.

Теорема 2. *Пусть* $\dim \mathscr{H} = \infty$. *Тогда множество ортогональных разложений единицы плотно в* $\mathfrak{M}(U)$.

Доказательство подобного результата содержится в работе Наймарка [1], где рассматривались спектральные функции симметричных операторов и использовалась слабая сходимость операторов. Приведем видоизменение этого доказательства для рассматриваемой нами ситуации.

Доказательство. Пусть $M \in \mathfrak{M}(U)$ и $\{P^{(n)}\}$ последовательность конечномерных проекторов, сильно сходящаяся к I в $\mathscr{H}$. Тогда $M^{(n)}(B) = P^{(n)}M(B)P^{(n)}$ $(B \in \mathscr{B}(U))$ — разложение единицы в конечномерном пространстве $\mathscr{H}^{(n)} = P^{(n)}\mathscr{H}$. Поскольку σ-алгебра $\mathscr{B}(U)$ является счетно-порожденной, то, согласно общей теореме Наймарка [1], $M^{(n)}$ можно продолжить до ортогонального разложения единицы $E^{(n)}$ в сепарабельном гильбертовом пространстве $\tilde{\mathscr{H}}^{(n)} \supseteq \mathscr{H}^{(n)}$. Так как $\dim \tilde{\mathscr{H}}^{(n)} \leqslant \dim \mathscr{H}$, то мы можем вложить $\tilde{\mathscr{H}}^{(n)}$ в $\mathscr{H}$ (так что $\mathscr{H}^{(n)}$ перейдет само в себя) и считать, что $\tilde{\mathscr{H}}^{(n)} \subseteq \mathscr{H}$. Тогда для любого о. п. S

$$\mu_S^{E^{(n)}}(B) - \mu_S^M(B) = 2\,\mathrm{Re}\,\mathrm{Tr}(I-P^{(n)})SP^{(n)}[E^{(n)}(B)-M(B)] + \\ + \mathrm{Tr}(I-P^{(n)})S(I-P^{(n)})[E^{(n)}(B)-M(B)],$$

откуда, учитывая что $\|E^{(n)}(B)-M(B)\| \leqslant 2$, получаем

$$\mathrm{var}\,(\mu_S^{E^{(n)}} - \mu_S^M) \leqslant 6\,\|(I-P^{(n)})S\|_1,$$

где $\|\cdot\|_1$ — ядерная норма (см., например, [2]). Но $\|(I-P^{(n)})S\|_1 \to 0$ для любого о. п. S. Это очевидно, если S — оператор конечного ранга; с другой стороны любой о. п. S аппроксимируется по ядерной норме операторами конечного ранга. Теорема доказана.

Доказательство обобщается на случай несепарабельного $\mathscr{H}$, с заменой последовательностей обобщенными последовательностями (сетями).

Из этой теоремы вытекает, что при естественных ограничениях, обеспечивающих непрерывность байесовского риска [7], в общей байесовской задаче с $\dim \mathscr{H} = \infty$ для любого $\varepsilon > 0$ существует ε-байесовское простое измерение.

ЛИТЕРАТУРА

[1] М. А. Наймарк, *Спектральные функции симметричного оператора*, Изв. Акад. наук СССР, сер. мат., 1 (1940), стр. 277-318.

[2] М. Рид и Б. Саймон, *Методы современной математической физики. I*, Мир, Москва 1977 (M. Reed and B. Simon, *Methods of modern mathematical physics. I*, Academic Press, 1972).

[3] К. Хелстром, *Квантовая теория проверки гипотез и оценивания*, Мир, Москва 1979 (C. Helstrom, *Quantum detection and estimation theory*, Academic Press, 1976).

[4] C. Helstrom, *Bayes-cost reduction algorithm in quantum hypotheses testing*, preprint, 1979.

[5] A. S. Holevo, *Statistical decision theory for quantum systems*, J. Multivariate Anal. 3 (1973), стр. 337-394.

[6] А. С. Холево, *Замечания об оптимальных квантовых измерениях*, Пробл. передачи информ. 10 (4) (1974), стр. 51-55.

[7] — *Исследования по общей теории статистических решений*, Наука, Москва, Тр. МИАН 124 (1976).

[8] H. Yuen, R. Kennedy and M. Lax, *Optimum testing of multiple hypotheses in quantum detection theory*, Proc. IEEE, IT-21, No. 2 (1975), p. 125-134.

Математический институт им. В. М. Стеклова
Академия наук СССР
Москва 117333, Вавилова 42, СССР

Received on 31. 3. 1980