Proceedings Fifteenth Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL 1977.

QUANTUM DETECTION AND ESTIMATION THEORY [(1)]

SANJOY K. MITTER[(2)] and S.K. YOUNG[(3)]

[(2)] Department of Electrical Engineering and Computer Science MIT, Cambridge, MA 02139

[(3)] Department of Mathematics, MIT, Cambridge, MA 02139 Now at BDM Corporation, McLean, VA

1. INTRODUCTION

In classical communication theory, the message to be transmitted is modulated and the resulting signal propagates through a given channel to produce a received waveform. The function of the receiver is to recover the signal from the received waveform, perhaps in an optimum manner, optimum being defined by some performance criterion. The input to the receiver may have some additive noise added to the received waveform. It is assumed that the receiver can be constructed independently of the model of the received waveform and the additive noise. Moreover, it is assumed that the optimum receiver can be physically realized.

In communication at optical frequencies neither of these two assumptions are valid. No matter what measurement we make of the received field, the outcome is random whose statistics depend on the measurement being made. This is a reflection of the laws of quantum physics. Furthermore, there is no guarantee that the measurement characterizing the receiver can be actually implemented.

In this paper, we announce some results on the M-any quantum detection problem. Full details will be published elsewhere [1]. For related work, see [2] and [3]. The proof in [2] appears to be incorrect and in [3], a complete duality theory is not presented.

It will be assumed that the reader is familiar with the notions of convex analysis in infinite dimensional spaces as for example, presented in [4].

In the classical formulation of detection theory (Bayesian hypothesis testing) it is desired to decide which of n possible hypotheses $H_1,\dots,H_n$ is true, based on observation of a random variable whose probability distribution depends on the several hypotheses. The decision entails certain costs that depend on which hypothesis is selected and which hypothesis corresponds to the true state of the system. A decision procedure or strategy prescribes which hypothesis is to be chosen for each possible outcome of the observed data; in general, it may be necessary to use a randomized strategy which specifies the probabilities with which each hypothesis should be chosen as a function of the observed data. The detection problem is to determine an optimal decision strategy.

In the quantum formulation of the detection problem, each hypothesis H_j corresponds to a possible state ρ_j of the quantum system under consideration. Unlike the classical situation, however, it is not possible to measure all relevant variables associated with the state of the system and to specify meaningful probability distributions for the resulting values. For the quantum detection problem it is necessary to specify not only the procedure for processing the experimental data, but also what data to measure in the first place. Hence the quantum detection problem involves determining the entire measurement process, or, in mathematical terms, determining the probability operator measure corresponding to the measurement process (for a discussion see [5]).

2. FORMULATION OF THE QUANTUM DETECTION PROBLEM

Let H be a separable complex Hilbert space corresponding to the physical variables of the system under consideration. There are n hypotheses $H_1,\dots,H_n$ about the state of the system, each corresponding to a different density operator ρ_j; every ρ_j is a nonnegative definite selfadjoint trace-class operator on H with trace 1 and is the analog of the distribution functions in the classical problem. Let S denote the set $S = \{1,\dots,n\}$. A general decision strategy is determined by a probability operator measure (POM) $m : 2^S \to L_s(H)_+$ [(1)]; in this case the POM effecting the decision needs only n components $m_1,\dots,m_n$ where each m_j is a positive selfadjoint bounded linear operator on H and

$$\sum_{i=1}^{n} m_i = I \quad . \tag{1}$$

The measurement outcome is an integer $i \in S$; the conditional probability that the hypothesis H_i is chosen when the state of the system is ρ_j is given by

$$\Pr\{i|j\} = \mathrm{tr}(\rho_j m_i) \qquad i,j = 1,\dots,m \quad . \tag{2}$$

We remark that it is crucial here to formulate the problem in terms of general probability operator measures rather than resolutions of the identity. For example, an instrument which simply chooses an arbitrary hypothesis with probability $1/n$ without even interacting with the system corresponds to a measurement process with the POM given by

$$m_j = I/n \quad ;$$

these are certainly not projections.

We denote by C_{ij} the cost associated with choosing hypothesis H_i when H_j is true. For a specified decision procedure effected by the POM $\{m_1,\dots,m_n\}$, the risk function is the conditional expected cost given that

(1) This research has been supported by NSF Grant #ENG76-02860.

(1) $L_s(H)_+$ denotes the space of bounded positive selfadjoint operators on H.

the system is in the state ρ_j , i.e.,

$$R_m(j) = tr[\rho_j \sum_{i=1}^{n} c_{ij} m_i] \quad .$$

If now μ_j specifies a prior probability for hypothesis H_j , the Bayes cost is the posterior expected cost

$$R_m = \sum_{j=1}^{n} R_m(j)\mu_j = tr \sum_{i=1}^{n} f_i m_i \tag{3}$$

where f_i is the selfadjoint trace-class operator

$$f_i = \sum_{j=1}^{n} c_{ij}\mu_j \rho_j \qquad i = 1,\ldots,n \quad . \tag{4}$$

The quantum detection problem is to find $m_1,\ldots,m_n$ so as to minimize (3) subject to the constraint (1) and subject to the condition that the operators m_j be selfadjoint and nonnegative definite, $m_j \geq 0$.

3. THE QUANTUM DETECTION PROBLEM AND ITS DUAL

Let H be a complex Hilbert space. The real linear space of compact self-adjoint operators $K_s(H)$ with the operator norm is a Banach space whose dual is isometrically isomorphic to the real Banach space $\tau_s(H)$ of self-adjoint trace-class operators with the trace norm, i.e., $K_s(H)^* = \tau_s(H)$ under the duality

$$\langle A,B\rangle = tr(AB) \leq |A|_{tr}|B| \qquad A\epsilon\tau_s(H),\ B\epsilon K_s(H) \ .$$

Here, $|B| = \sup\{|B\phi| : \phi\epsilon H, |\phi| \leq 1\} = \sup\{tr(AB) : A\epsilon\tau_s(H), |A|_{tr} \leq 1\}$ and $|A|_{tr}$ is the trace norm $\sum_i |\lambda| < +\infty$ where $A\epsilon\tau_s(H)$ and $\{\lambda_i\}$ are the eigenvalues of A repeated according to multiplicity. The dual of $\tau_s(H)$ with the trace norm is isometrically isomorphic to the space of all linear bounded self-adjoint operators, i.e. $\tau_s(H)^* = L_s(H)$ under the duality

$$\langle A,B\rangle = tr(AB) \qquad A\epsilon\tau_s(H),\ B\epsilon L_s(H).$$

Moreover the orderings are compatible in the following sense. If $K_s(H)_+$, $\tau_s(H)_+$, and $L_s(H)_+$ denote the closed convex cones of nonnegative definite operators in $K_s(H)$, $\tau_s(H)$, and $L_s(H)$ respectively, then

$$[K_s(H)_+]^* = \tau_s(H)_+ \quad \text{and} \quad [\tau_s(H)_+]^* = L_s(H)_+$$

where the associated dual spaces are to be understood in the sense defined above.

Let f_j be given elements of $\tau_s(H)$ (as defined in (4)), $j=1,\ldots n$. Define the functionals $F_j: L_s(H) \to \overline{R}$ by

$$F_j(A) = \delta_{\geq o}(A) + tr(f_j A) \qquad A\epsilon L_s(H),\ j = 1,\ldots,n. \tag{5}$$

where $\delta_{\geq o}(\cdot)$ denotes the indicator function for the set $L_s(H)_+$ of non-negative definite operators, i.e. $\delta_{\geq o}(A)$ is 0 if $A \geq 0$ and $+\infty$ otherwise. Each F_j is proper convex and w*-lower semicontinuous on $L_s(H)$, since $L_s(H)_+$ is a w*-closed convex cone and $A \mapsto tr(f_j A)$ is a continuous (in fact w*-continuous) linear functional on $L_s(H)$. Define the function $G: L_s(H) \to \overline{R}$ by

$$G(A) = \delta_{\{o\}}(A), \qquad A\epsilon L_s(H), \tag{6}$$

that is $G(A)$ is 0 if $A = 0$ and $G(A)$ is $+\infty$ if $A \neq 0$; G is trivially convex and lower semicontinuous. Let $m = (m_1,\ldots,m_n)$ denote an element of $L_s(H)^n$, the Cartesian product of n copies of $L_s(H)$. Then the quantum detection problem (3) may be written

$$P = \inf\left\{\sum_{j=1}^{n} F_j(m_j) + G(I-Lm) : m = (m_1,\ldots,m_n) \ \epsilon\ L_s(H)^n\right\} \tag{7}$$

where $L: L_s(H)^n \to L_s(H)$ is the continuous linear operator

$$L(m) = \sum_{j=1}^{n} m_j, \quad m\ \epsilon\ L_s(H)^n. \tag{8}$$

We consider a family of perturbed problems defined by

$$P(A) = \inf\left\{\sum_{j=1}^{n} F_j(m_j) + G(A-Lm) : m\epsilon L_s(H)^n\right\}, \quad A\epsilon L_s(H). \tag{9}$$

$P(\cdot)$ is a convex function $L_s(H) \to \overline{R}$ and $P = P(I)$. Note that we are taking perturbations in the equality constraint, i.e. the problem $P(A)$ requires that every feasible m satisfy $Lm = A$. We remark that $G(\cdot)$ is nowhere continuous, so that there is certainly no Kuhn-Tucker point $\overline{m}$ such that $G(\cdot)$ is continuous at $L\overline{m}$.

In order to construct the dual problem corresponding to the family of perturbed problems (9) we must calculate the conjugate functions of F_j and G. We would like to pose the dual problem in the space $\tau_s(H)$, so we consider $L_s(H) = \tau_s(H)^*$ and compute the pre-conjugates of F_j,G. It can be shown that the dual problem is

$$(^*P)(I) = \sup\{tr(y) : y\epsilon\tau_s(H),\ f_j - y \geq 0 \qquad j = 1,\ldots,n\}. \tag{10}$$

It is not too difficult to show that the primal problem has a solution and that there is no duality gap. The difficult part is to show that the dual problem has solutions. It turns out that the level sets of the dual cost functions are bounded in $\tau_s(H)$ but not weakly compact. This suggests that we imbed $\tau_s(H)$ in its bidual $\tau_s(H)^{**} = L_s(H)^*$ and extend the dual problem to the larger space; it will then turn out that there are solutions in $\tau_s(H)$. This approach works because $\tau_s(H)$ has a natural topological complement as a subset of $L_s(H)^*$.

Proposition 1. $L_s(H)^* = \tau_s(H) \oplus_1 (JK_s(H))$ where J is the canonical imbedding of $K_s(H)$ in $L_s(H)$. In other words, every bounded linear functional y on $L_s(H)$ may be uniquely represented in the form $y = y_{ac} \oplus y_{sg}$ where $y_{ac} \in \tau_s(H)$ and $y_{sg} \in K_s(H)$, and

$$y(A) = tr(y_{ac}A) + y_{sg}(A),\ A \in L_s(H)$$

$$|y| = |y_{ac}|_{tr} + |y_{sg}|. \blacksquare$$

Before calculating the dual problem, it is necessary to determine what the positive linear functions look like in terms of the decomposition provided by Proposition 1.

Proposition 2. Let $y \in L_s(H)^*$. Then $y \in [L_s(H)_+]^+$ iff $y_{ac} \in \tau_s(H)_+$ and $y_{sg} \in [L_s(H)_+]^+$. ■

It can be shown that in this enlarged space the dual problem $*(P^*)(I) = \sup_y[y(I)-P^*(y)]$ is given by

$$*(P^*)(I) = \sup\{tr(y_{ac})+y_{sg}(I) : y \in L_s(H)^*,\ y_{sg} \leq 0,\ y_{ac} \leq f_j,\ j = 1,\ldots,n\}.$$

Proposition 3. $P(\cdot)$ is continuous at I, and hence $\partial P(I) \neq \emptyset$. In particular, $*(P^*)(I) = P(I)$ and the dual problem $*(P^*)(I)$ has solutions. ■

It is now an easy matter to show that the dual problem actually has solutions in $\tau_s(H)$, that is solutions in $L_s(H)^*$ with 0 singular part. This leads to the main theorem.

MAIN THEOREM. Let H be a complex Hilbert space and suppose $(f_1,\ldots,f_n) \in \tau_s(H)^n$. Then the quantum detection problem has solutions. Moreover, the following statements are equivalent for $m=(m_1,\ldots,m_n) \in L_s(H)^n$:

1) m solves the quantum detection problem

2) $\sum_{j=1}^{n} m_j = I;\ m_i \geq 0$ for $i=1,\ldots,n$;

$\sum_{j=1}^{n} f_j m_j \leq f_i$ for $i=1,..,n$

3) $\sum_{j=1}^{n} m_j = I;\ m_i \geq 0$ for $i=1,\ldots,n$;

$\sum_{j=1}^{n} m_j f_j \leq f_i$ for $i=1,\ldots,n$.

Under any of the above conditions if follows that

$y = \sum_{j=1}^{n} f_j m_j = \sum_{j=1}^{n} m_j f_j$ is self-adjoint and is the unique solution of the dual problem.

4. FINAL REMARKS

Quantum estimation theory is the continuous counterpart of the problem treated here. The main difficulties here are in problem formulation and having an adequate theory of integration with respect to operator-valued measures. For details, see [1] and [6].

REFERENCES

[1] S.K. Mitter and S.K. Young: Quantum Detection and Estimation Theory, I,II, to appear.

[2] H. Yuen, R. Kennedy and M. Lax: Optimum Testing of Multiple Hypotheses in Quantum Detection Theory, IEEE Trans. on Information Theory, IT-21, 1975, p. 25.

[3] A.S. Holevo: Remarks on Optimal Quantum Measurements, Problemy Peredachi Informatsii, Vol. 10, No. 4, 1974, p. 51.

[4] I. Ekeland and R. Temam: Analyse Convexe et Problèmes Variationnels, Dunod-Gauthier-Villars, 1974.

[5] E.B. Davies: Quantum Theory of Open Systems, Academic Press, New York, 1976.

[6] S.K. Young: Doctoral dissertation, M.I.T., August 1977.