ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2022, Vol. 43, No. 7, pp. 1646–1650. © Pleiades Publishing, Ltd., 2022.

On Optimization Problem for Positive Operator-Valued Measures

A. S. Holevo*

(Submitted by A. M. Elizarov)

Steklov Mathematical Institute of Russian Academy of Sciences, Moscow, 119991 Russia

Received June 4, 2022; revised June 24, 2022; accepted June 28, 2022

Abstract—It is shown that a number of optimization problems in quantum information theory: the χ-capacity (called the Holevo capacity in literature) of a quantum channel; the classical capacity of quantum observable; entanglement of formation—can be recast as a generalization of a Bayes problem over the set of all quantum states. This allows us to consider it as a convex programming problem for which necessary and sufficient optimality conditions along with the dual problem can be formulated.

DOI: 10.1134/S1995080222100158

Keywords and phrases: *quantum state ensemble, positive operator-valued measure, convex optimization.*

Let $\mathcal{H}$ be a finite-dimensional Hilbert space, $\mathfrak{T}(\mathcal{H})$ the Banach space of operators on $\mathcal{H}$ equipped with trace norm, $\mathfrak{S} = \mathfrak{S}(\mathcal{H})$ the compact convex set of quantum states equipped with trace-norm distance. An *ensemble* is a probability measure $\pi(dS)$ on the set $\mathfrak{S}$, cf. [2], and the set of all such measures is denoted $\mathcal{P}(\mathfrak{S})$.

Let f be a continuous concave function on the compact convex set $\mathfrak{S}$. The proof of theorem below uses very little of other special properties of f and $\mathfrak{S}$. Consider the functional

$$F(\pi) = \int_{\mathfrak{S}} f(S)\pi(dS), \tag{1}$$

on $\mathcal{P}(\mathfrak{S})$. From the definition, it is a continuous affine functional. We are interested in minimization of this functional on the closed convex subset $\mathcal{P}_{\overline{S}}$ of probability measures π with the fixed given barycenter

$$\overline{S} = \int_{\mathfrak{S}} S\,\pi(dS). \tag{2}$$

Under mild additional conditions the functional $F(\pi)$ attains its minimum $\mathcal{F}(\overline{S})$ on the compact set $\mathcal{P}_{\overline{S}}$. The resulting function $\mathcal{F}(S), S \in \mathfrak{S}$, is convex, in fact it is equal to the *convex closure* of $f(S)$ i.e. the greatest lower semicontinuous convex function majorized by $f(S)$ [7]. Concavity of f implies then that we can choose the minimizing measure π to be supported by the pure states since we can always make spectral decompositions of all the density operators S into pure states without changing the barycenter and without increasing the value $F(\pi)$.

This problem is relevant to a number of issues in quantum information theory.

1. Computation of the χ-capacity [2] of a quantum channel Φ defined as

$$C_\chi(\Phi) = \sup_\pi \left[H\left(\Phi\left[\int_{\mathfrak{S}} S\,\pi(dS)\right]\right) - \int_{\mathfrak{S}} H(\Phi[S])\pi(dS)\right]$$

*E-mail: holevo@mi-ras.ru

$$= \sup_{\overline{S}} \left[H\left(\Phi\left[\overline{S}\right]\right) - \inf_{\pi\in\mathcal{P}_{\overline{S}}} \int\limits_{\mathfrak{S}} H(\Phi[S])\pi(dS) \right].$$

Here the second term in the squared brackets is the convex closure of the channel output entropy, with

$$f(S) = H(\Phi[S]) = \mathrm{Tr}\, K(S)S,$$
$$K(S) = -\Phi^{*}\left[\log \Phi[S]\right].$$

2. A similar case is the classical capacity of quantum-classical channel (observable) given by the map $M : S \to p_S(z) = \mathrm{Tr}\, Sm(z)$, where $m(z)$ is a uniformly bounded positive-operator-valued function of $z \in Z$, such that $\int m(z)dz = I$ (the unit operator). Here (Z, dz) is a measure space. Then $p_S(z)$ is the probability density of the outcomes of the quantum measurement described by $m(z)$. In this case the classical capacity of the channel is equal to the χ-capacity (the channel is entanglement-breaking) and

$$C(M) = C_{\chi}(M) = \sup_{S} \left[h\left(p_S\right) - \inf_{\pi:\overline{S}_{\pi}=S} \int\limits_{\mathfrak{S}} h\left(p_S\right)\pi(dS) \right],$$

where $h\left(p_S\right) = -\int p_S(z)\log p_S(z)dz$ is the differential entropy of this probability density and

$$f(S) = h\left(p_S\right) = \mathrm{Tr}\ K(S)S,$$
$$K(S) = -\int m(z)\log p_S(z)dz.$$

3. Another case of interest described e.g. in Sec. 7.5 of [1] is the Entanglement of Formation. Let S_{12} be a state in the tensor product of two Hilbert spaces $\mathcal{H}_1 \otimes \mathcal{H}_2$. *Entanglement of Formation* of the state S_{12} is the convex closure of $H(S_1)$, where $S_1 = \mathrm{Tr}_2 S_{12}$ is the partial state:

$$E_F(S_{12}) = \inf_{\pi:\overline{S}_{\pi}=S_{12}} \int\limits_{\mathfrak{S}(\mathcal{H}_1\otimes\mathcal{H}_2)} H(\mathrm{Tr}_2 S)\,\pi(dS),$$

minimization is over all probability measures $\pi(dS)$ on $\mathfrak{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ satisfying $\int_{\mathfrak{S}(\mathcal{H}_1\otimes\mathcal{H}_2)} S\,\pi(dS) = S_{12}$. In that case $f(S_{12}) = H(S_1), K(S_{12}) = -\left(\log S_1 \otimes I_2\right)$.

In all these cases we are looking for the solution of the convex programming problem

$$F(\pi) \equiv \int\limits_{\mathfrak{S}} \mathrm{Tr}\ K(S)S\,\pi(dS) \longrightarrow \min$$
$$\pi \in \mathcal{P}(\mathfrak{S})$$
$$\int\limits_{\mathfrak{S}} S\,\pi(dS) = \overline{S}, \tag{3}$$

where $\overline{S}$ is a fixed density operator. We will give the duality relation and necessary and sufficient conditions for optimality basing on the results obtained in the monograph [5].

We start by introducing an equivalent but more convenient definition: we now call *ensemble* a measure $\Pi(dS)$ on $\mathfrak{S}$ with values in the positive cone of $\mathfrak{T}(\mathcal{H})$, such that $\Pi(\mathfrak{S}) \in \mathfrak{S}$. We call $\Pi(\mathfrak{S}) \equiv \overline{S}_{\Pi}$ the *average state* of the ensemble. The equivalence with the initial definition is established by the relations

$$\Pi(dS) = S\pi(dS); \quad \overline{S}_{\Pi} = \int\limits_{\mathfrak{S}} S\pi(dS),$$

where $\pi(A) = \mathrm{Tr}\,\Pi(A)$, A is any Borel subset of $\mathfrak{S}$. Then the minimized functional can be rewritten as the *scalar integral* of the operator-valued function $K(S)$ with respect to the operator-valued measure

$\Pi(dS)$, the construction of which was given in the infinite-dimensional case in Ch. I of [5] (see also [4]):

$$F(\pi) = \int_{\mathfrak{S}} \langle K(S), \Pi(dS) \rangle \, .$$

In what follows we assume that $\mathcal{H}$ *is finite-dimensional*. In that case case we can just assume that $K(S)$ is measurable function with values in the cone of positive operators on $\mathcal{H}$ and the scalar integral can be understood via the expression $\int_{\mathfrak{S}} \mathrm{Tr}\; K(S)S\, \pi(dS)$. Then the optimization problem

$$\int_{\mathfrak{S}} \langle K(S), \Pi(dS) \rangle \longrightarrow \min$$
$$\Pi(A) \geq 0 \text{ for any Borel } A \subseteq \mathfrak{S},$$
$$\Pi(\mathfrak{S}) = \overline{S},$$

becomes similar to a generalization of the Bayes problem [4, 5]. Combination of Theorem 2.1 and Theorem 2.2 from Ch. II of [5] implies

Theorem. *The problem dual to (3) is*

$$\max\left\{ \mathrm{Tr}\, \overline{S}\Lambda : \Lambda^* = \Lambda,\, \Lambda \leq K(S) \text{ for all } S \in \mathfrak{S} \right\} . \tag{4}$$

The following statements are equivalent:

(i) $\Pi_0(dS)$ *is the solution of the problem* (3); Λ_0 *is the solution of the problem* (4);

(ii) **a.** $\Lambda_0 \leq K(S)$ *for all* $S \in \mathfrak{S}$;

b. $\int_A [K(S) - \Lambda_0]\; \Pi_0(dS) = 0$ *for any Borel subset* $A \subseteq \mathfrak{S}$.

The condition (ii.b) can be rewritten as

$$[K(S) - \Lambda_0]\; S = 0 \quad (\mathrm{mod}\, \pi_0),$$

which means that the equality holds a.e. with respect to the measure π_0. By integrating, we obtain

$$\int_{\mathfrak{S}} K(S)\, S\, \pi_0(dS) = \Lambda_0 \overline{S},$$

which gives equation for determination of Λ_0. Note that Λ_0 must be Hermitean operator.

In the case of measurement channel, this equation reduces to

$$-\int_{\mathfrak{S}} \int m(z) \log p_S(z) dz\, S\, \pi_0(dS) = \Lambda_0 \overline{S}.$$

For completeness, we give the proof of the Theorem, taking into account simplifications due to finite dimensionality of $\mathcal{H}$. An infinite-dimensional generalization of the Theorem would have important applications to ensemble optimization problems.

Proof. In what follows we use the notation $\langle X, Y \rangle = \mathrm{Tr}\, X\, Y$. Let us fix $S_0 \in \mathfrak{S}$ and show that

$$\inf\left\{ \int_{\mathfrak{S}} \langle K(S), \Pi(dS) \rangle : \Pi(\mathfrak{S}) = \overline{S} \right\}$$
$$= \inf\left\{ \int_{\mathfrak{S}} \langle K(S) - K(S_0), \Pi(dS) \rangle : \Pi(\mathfrak{S}) \leq \overline{S} \right\} + \left\langle K(S_0), \overline{S} \right\rangle .$$

It is sufficient to show that

$$\inf\left\{\int\limits_{\mathfrak{S}} \langle K(S)-K(S_0),\Pi(dS)\rangle : \Pi(\mathfrak{S})=\overline{S}\right\}$$
$$=\inf\left\{\int\limits_{\mathfrak{S}} \langle K(S)-K(S_0),\Pi(dS)\rangle : \Pi(\mathfrak{S})\leq\overline{S}\right\}. \tag{5}$$

Let Π be such that $\Pi(\mathfrak{S})\leq\overline{S}$. Defining

$$\overline{\Pi}(A)=\Pi(A)+\left(\overline{S}-\Pi(\mathfrak{S})\right)1_A(S_0),$$

we get $\overline{\Pi}(\mathfrak{S})=\overline{S}$ and equality in (5).

Denote $G(\Pi)=\Pi(\mathfrak{S})-\overline{S}$ and consider the problem of minimizing the functional

$$F(\Pi)=\int\limits_{\mathfrak{S}} \langle K(S)-K(S_0),\Pi(dS)\rangle+\left\langle K(S_0),\overline{S}\right\rangle$$

over the convex set of Π's satisfying $G(\Pi)\leq 0$. For this we compute the dual functional

$$\varphi(S)=\inf\left\{F(\Pi)+\langle S,G(\Pi)\rangle\right\},$$

where the infimum is over the set of all positive $\mathfrak{S}$−valued measures. We show that

$$\varphi(S)=\begin{cases}\left\langle K(S_0),\overline{S}\right\rangle-\left\langle S,\overline{S}\right\rangle, & \text{if } S\geq K(S_0)-K(S'), \quad \forall S';\\ -\infty & \text{otherwise.}\end{cases} \tag{6}$$

Let S be such that for some S' and some $X\geq 0$

$$\langle S,X\rangle<\left\langle K(S_0)-K(S'),X\right\rangle.$$

Defining $\Pi_n(A)=nX\,1_A(S')$, we have

$$\varphi(S)\leq n\left[\langle S,X\rangle-\left\langle K(S_0)-K(S'),X\right\rangle\right]+\left\langle K(S_0),\overline{S}\right\rangle-\left\langle S,\overline{S}\right\rangle$$

whence $\varphi(S)=-\infty$.

Let now

$$S\geq K(S_0)-K(S') \quad \forall S'. \tag{7}$$

By letting $\Pi(A)\equiv 0$, we obtain $\varphi(S)\leq\left\langle K(S_0),\overline{S}\right\rangle-\left\langle S,\overline{S}\right\rangle$. The converse inequality follows from

$$\langle S,\Pi(\mathfrak{S})\rangle\geq\int\limits_{\mathfrak{S}}\left\langle K(S_0)-K(S'),\Pi(dS')\right\rangle, \tag{8}$$

which is obtained from (7) by integration (see Lemma 2.1 in Ch. I of [5]). This implies the first line in (6).

According to the general Lagrange duality theorem (see Appendix)

$$\inf_{G(\Pi)\leq 0}F(\Pi)=\max\left\{\varphi(S):S\geq 0\right\},$$

i.e. taking into account (5), (6)

$$\inf\left\{\int\limits_{\mathfrak{S}} \langle K(S),\Pi(dS)\rangle : \Pi(\mathfrak{S})=\overline{S}\right\}$$
$$=\max\left\{\left\langle K(S_0),\overline{S}\right\rangle-\left\langle S,\overline{S}\right\rangle: \quad S\geq K(S_0)-K(S') \quad \forall S'\right\}.$$

Denoting $\Lambda=K(S_0)-S$, we come to (4).

Let (i) be fulfilled, (4) implies

$$\int\limits_{\mathfrak{S}} \langle K(S), \Pi_0(dS)\rangle = \left\langle \Lambda_0, \overline{S}\right\rangle. \tag{9}$$

The inequality (iia): $\Lambda_0 \leq K(S)$, $S \in \mathfrak{S}$ holds by (4). It follows for any Borel $A \subseteq \mathfrak{S}$:

$$\langle \Lambda_0, \Pi_0(A)\rangle = \int\limits_{A} \langle \Lambda_0, \Pi_0(dS)\rangle \leq \int\limits_{A} \langle K(S), \Pi_0(dS)\rangle. \tag{10}$$

By (9) it should be equality here, i.e. $\int_A \langle K(S) - \Lambda_0, \Pi_0(dS)\rangle = 0$. But since $K(S) - \Lambda_0 \geq 0$, this implies (iib) (for detail see Proposition 3.3 from Ch. I of [5]).

Conversely, for arbitrary Π satisfying $\Pi(\mathfrak{S}) = \overline{S}$ we have by (iia)

$$\int\limits_{\mathfrak{S}} \langle K(S), \Pi(dS)\rangle \geq \left\langle \Lambda_0, \overline{S}\right\rangle$$

and taking $A = \mathfrak{S}$ in (iib) we obtain (9) whence (i) follows. □

Appendix. Let F be a convex functional on a convex subset $\mathfrak{S}$ of a linear subspace L and G be a convex map of $\mathfrak{S}$ into partially ordered Banach space L_1.

Consider the optimization problem

$$F(x) \longrightarrow \inf, \quad x \in \mathfrak{S}; \quad G(x) \leq 0. \tag{11}$$

The following duality theorem holds (see e.g. [6], pp. 217, 224):

Theorem. *Assume that the positive cone of* L_1 *contains an inner point, and there exists* $x_1 \in S$ *such that* $\langle \lambda, G(x_1)\rangle < 0$ *for all* $\lambda \in L_1^*, \lambda > 0$. *Then if the quantity (11) is finite,*

$$\inf\{F(x): \quad x \in \mathfrak{S}; \quad G(x) \leq 0\} = \max\{\varphi(\lambda): \quad \lambda \in L_1^*, \quad \lambda \geq 0\}, \tag{12}$$

where

$$\varphi(\lambda) = \inf_{x \in \mathfrak{S}} \{F(x) + \langle \lambda, G(x)\rangle\}$$

is the dual functional.

Let λ_0 *be a solution of the dual problem in the right-hand side of (12). If the infimum in the left-hand side of (12) is attained on* x_0, *then* x_0 *is a solution of the problem* $\min\{F(x) + \langle \lambda_0, G(x)\rangle; x \in \mathfrak{S}\}$ *and* $\langle \lambda_0, G(x_0)\rangle = 0$.

FUNDING

This work is supported by Russian Science Foundation under the grant no. 19-11-00086, https://rscf.ru/project/19-11-00086/. The author is grateful to M. E. Shirokov for useful remarks.

REFERENCES

1. A. S. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction,* 2nd ed. (De Gruyter, Berlin, 2019).
2. A. S. Holevo and M. E. Shirokov, "Continuous ensembles and the capacity of infinite-dimensional quantum channels," Theory Probab. Appl. **50**, 86–98 (2005).
3. A. S. Holevo, "Statistical decision theory for quantum systems," J. Multivariate Anal. **3**, 337–394 (1973).
4. A. S. Holevo, "On a vector-valued integral in the noncommutative statistical decision theory," J. Multivariate Anal. **5**, 462–465 (1975).
5. A. S. Holevo, "Studies in general theory of statistical decisions," Proc. Steklov Math. Inst. **124** (1978).
6. D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1969).
7. M. E. Shirokov, "On properties of the space of quantum states and their application to construction of entanglement monotones," Izv.: Math. **74**, 849–882 (2010).