

# 測度空間における凸最適化

## — 無限次元における離散と連続 —

伊藤 聡<sup>†</sup>

(受付 2012 年 12 月 12 日; 改訂 2013 年 6 月 20 日; 採択 7 月 1 日)

### 要 旨

すべての最適化問題は測度に関する最適化問題である。とは言い過ぎであるとしても、無限次元の最適化問題は、ほとんどの場合、適当な測度の空間における最適化問題であるとみなすことができる。一般に測度は Lebesgue 測度に関して絶対連続な成分と離散的な成分に分解することができるが、測度空間上の最適化において、その解が離散測度になることは少なくない。本稿は、どのような状況のもとで最適解が離散測度もしくは絶対連続測度になるのかを明らかにする試みの端緒として、まずコンパクトな空間上の測度の凸最適化についてまとめてみたものであり、公益社団法人日本オペレーションズ・リサーチ学会の第 24 回 RAMP (数理計画研究部会) シンポジウムにおける講演予稿に加筆訂正を加えたものである。

キーワード：確率測度，無限計画，最適制御，半無限計画，一般化モーメント問題，通信路容量。

### 1. はじめに

無限次元の最適化問題は、多くの場合、適当な測度の空間における最適化問題とみなすことができる。一般に測度は Lebesgue 測度に関して絶対連続な成分と離散的な成分に分解することができるが、測度空間上の最適化において、解が離散測度になる状況が予想外に多いのは何故なのか、あるいは逆にどのような状況のもとで解が絶対連続測度となるのだろうか。本稿においては、コンパクトな空間上の測度の凸最適化に限定して考えてみたい。

本稿の構成は以下の通りである。まず第 2 節で、測度が双対変数として現れる問題として最適制御問題と半無限計画問題を取り上げる。次に、測度が主変数として現れる問題として、第 3 節では容量問題とモーメント問題の一般化として定式化される線形最適化、また離散測度の極限としてこれらの最適解を求める反復法、そして確率測度の線形最適化について述べる。最後に、第 4 節で測度空間上の非線形凸最適化および情報理論における通信路容量問題との関連について触れる。

$X$  をコンパクトな Hausdorff 空間とする。 $X$  の Borel 集合族  $\mathcal{B}$  上の有限な測度  $\mu$  (すなわち  $\mu(X) < \infty$ ) を  $X$  上の有限 Borel 測度という。 $X$  上の 2 つの有限 Borel 測度の差として表される符号つき測度 (これを  $X$  上の有限な符号つき Borel 測度という) からなる空間を  $M(X)$  と書く。自然な線形演算により測度空間  $M(X)$  は線形空間となるが、 $X$  上の有限 Borel 測度から

---

<sup>†</sup> 統計数理研究所：〒 190-8562 東京都立川市緑町 10-3

なる部分集合  $M(X)_+$  はこのとき凸錐となる. したがって, 半順序  $\geq$  を

$$\mu \geq 0 := \mu \in M(X)_+$$

と定める(このとき  $M(X)_+$  は非負錐と呼ばれる) ことにより  $M(X)$  は順序線形空間となる. またいわゆる全変動をノルムと定めると  $M(X)$  は完備すなわち Banach 空間となる. 特に  $X \subset \mathbb{R}$  のとき  $M(X)$  は Stieltjes 測度の空間となり, これは  $X$  上の 2 つの非減少関数の差として表される有界変動関数の空間  $BV(X)$  と同一視できる.

一方,  $X$  上のすべての連続関数からなる線形空間に一樣ノルムを定義することにより得られる Banach 空間を  $C(X)$  とすると,  $M(X)$  は  $C(X)$  の双対空間  $C(X)^*$  とみなせる (Riesz もしくは Riesz–Markov–角谷の表現定理, 例えば田辺, 1978 や Yosida, 1980 を参照). また, このとき  $M(X)$  の非負錐  $M(X)_+$  は  $C(X)$  の非負錐  $C(X)_+$  の双対錐  $C(X)_+^*$  すなわち

$$\mu \geq 0 \iff \int_X f(x) d\mu \geq 0 \quad \forall f \geq 0$$

とみなせる.

以下では,  $X$  上の符号つき測度  $\mu \in M(X)$  は,  $X \subset \mathbb{R}$  のとき, あるいはどの空間で定義された測度なのかを明確にしたいとき,  $\mu(x)$  などと書くことがある.

## 2. 双対変数としての測度の最適化

測度の最適化問題として最初に挙げるべきものは, 測度が双対変数として現れる問題ではないだろうか. 本節では, そのような問題の例として最適制御問題と半無限計画問題を見てみよう. 両者の共通点は不等式制約が関数空間上で定義されていることである. 以下では, 関数空間のベクトルとみたとき, 関数を単に変数と表現することがあるので注意されたい.

### 2.1 最適制御

連続時間の開ループ最適制御問題の例として, 不等式制約条件を持つ Bolza 型の非線形最適制御問題

$$(2.1) \quad \begin{cases} \min_u g_0(x(T), T) + \int_0^T f_0(x(t), u(t), t) dt \\ \text{subject to } \dot{x}(t) = f(x(t), u(t), t) \quad \forall t \in [0, T], \quad x(0) = x_0 \\ g(x(t), u(t), t) \leq 0 \quad \forall t \in [0, T] \end{cases}$$

を考える. 制御系の状態を  $x(t) \in \mathbb{R}^n$ , 入力を  $u(t) \in \mathbb{R}^r$  とし, 関数  $g_0: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $f_0: \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $f: \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R} \rightarrow \mathbb{R}^n$ ,  $g: \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R} \rightarrow \mathbb{R}^m$  の滑らかさは適当に仮定するものとする. 終端時刻  $T$  は有限に固定されていても, あるいは自由であってもよい. また, ここでは簡単のため  $m=1$  とする.

この最適制御系に対する随伴系を次のように表現することができる (Girsanov, 1972; Ito and Shimizu, 1990).

$$(2.2) \quad \begin{aligned} \psi(t) = & \nabla_x g_0(x(T), T) + \int_t^T \{ \nabla_x f_0(x(\tau), u(\tau), \tau) \\ & + \nabla_x f(x(\tau), u(\tau), \tau) \psi(\tau) \} d\tau + \int_t^T \nabla_x g(x(\tau), u(\tau), \tau) d\lambda(\tau) \end{aligned}$$

これは, 例えば  $x \in C([0, T], \mathbb{R}^n)$ ,  $u \in C([0, T], \mathbb{R}^r)$ ,  $\psi \in BV([0, T], \mathbb{R}^n)$ ,  $\lambda \in BV([0, T], \mathbb{R})$  として  $(C(X, Y), BV(X, Y))$  はそれぞれ  $X$  から  $Y$  への連続関数もしくは有界変動関数からなる Banach

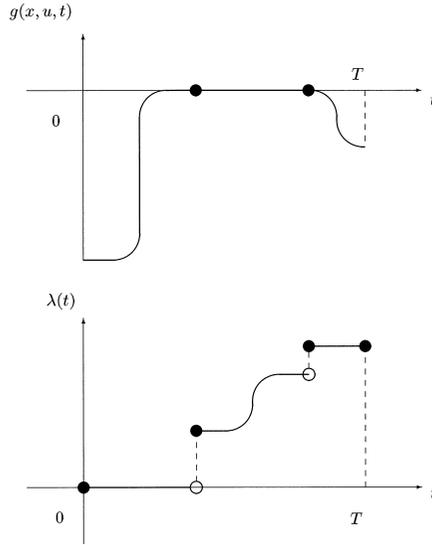


図 1. 不等式制約の接合点における乗数  $\lambda$  の跳躍.

空間), Lagrange 関数  $L$  を

$$L(x, u, \psi, \lambda) := g_0(x(T), T) + \int_0^T f_0(x(t), u(t), t) dt + \psi(t)^T \left\{ x(t) - x_0 - \int_0^t f(x(\tau), u(\tau), \tau) d\tau \right\} + \int_0^T g(x(t), u(t), t) d\lambda(t)$$

と定義して得られる Karush–Kuhn–Tucker 条件 (KKT 条件) の一部である (Ito and Shimizu, 1990 の (7) 式).

したがって  $\psi$  は状態  $x$  に対する随伴変数を表すが, 問題 (2.1) を状態  $x$  と入力  $u$  に関する最適化問題とみた場合, 随伴変数  $\psi$  は等式制約条件となる状態方程式系に対する Lagrange 乗数であり, このとき随伴系 (2.2) は状態  $x$  に関する停留条件に相当する. また,  $\lambda$  は不等式制約条件に対する Lagrange 乗数である.

さて, 随伴系 (2.2) において,  $BV([0, T], \mathbb{R})$  は  $C([0, T], \mathbb{R}) = C([0, T])$  の双対空間  $M([0, T])$  と同一視できるので, 乗数  $\lambda$  は有界変動関数と見ることもできるし, Stieltjes 測度と見ることもできる. 図 1 に示すように  $\lambda$  が制約の接合点 (制約境界に触れるあるいは境界から離れる点) において跳躍し得ること, したがって随伴系 (2.2) により  $\psi$  もこれらの点において不連続になり得ることに注意されたい. なお, 図 1 では乗数  $\lambda$  が右連続な非減少関数として表現されているが, 両端点を除けば不連続点  $t_i$  における値  $\lambda(t_i)$  そのものは本質的ではなく, 最適性条件においては跳躍量  $\lambda(t_i + 0) - \lambda(t_i - 0)$  (ただし両端ではそれぞれ  $\lambda(0 + 0) - \lambda(0)$ ,  $\lambda(T) - \lambda(T - 0)$ ) のみが意味を持つ.

しかし, ある種の一次独立制約想定のもとでは, 乗数  $\lambda$  が Lebesgue 測度に関して絶対連続であること, したがって (2.2) の右辺は  $t$  に関して連続となり随伴変数  $\psi$  が連続であることが保証される.

定理 1. 各  $t \in [0, T]$  において

$$\nabla_u g_i(x(t), u(t), t), \quad i \in I(t)$$

は一次独立であるとする(ただし

$$I(t) := \{i \mid g_i(x(t), u(t), t) = 0\}$$

は  $t$  において活性(active)な制約式の添字集合). このとき, 不等式制約条件に対する乗数  $\lambda$  は Lebesgue 測度に関して絶対連続である.

証明. Berkovitz (1961) もしくは Ito and Shimizu (1990) を参照.  $\square$

## 2.2 半無限計画

半無限計画問題は一般に次のように定式化される.

$$(2.3) \quad \begin{cases} \min_{x \in \mathbb{R}^n} f(x) \\ \text{subject to } g(x, y) \leq 0 \quad \forall y \in Y \end{cases}$$

ここでは,  $Y$  をコンパクトな Hausdorff 空間とし,  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  は  $\mathbb{R}^n$  上で微分可能,  $g: \mathbb{R}^n \times Y \rightarrow \mathbb{R}^m$  は  $y$  に関して連続かつ  $\mathbb{R}^n \times Y$  上で  $x$  に関して連続微分可能であるとする. また, 簡単のため  $m=1$  とする.

Lagrange 汎関数  $L: \mathbb{R}^n \times M(Y) \rightarrow \mathbb{R}$  を

$$L(x, \lambda) := f(x) + \int_Y g(x, y) d\lambda$$

と定義すると, Karush-Kuhn-Tucker (KKT) 条件は

$$(2.4) \quad \begin{cases} \nabla_x L(x, \lambda) = \nabla f(x) + \int_Y \nabla_x g(x, y) d\lambda = 0 \\ \int_Y g(x, y) d\lambda = 0 \\ \lambda \geq 0 \end{cases}$$

となる.

問題 (2.3) の最適解に対して KKT 条件を満たす乗数  $\lambda$  が存在することは, 例えば Slater 型の制約想定のもとで保証されるが, 一方では, KKT 条件を満たす乗数はいつでも離散測度と考えてよいことが示される (伊藤, 1998; Ito et al., 2000).

定理 2. KKT 条件 (2.4) を満たす乗数の集合は, 空集合でなければ, 高々  $n$  個のサポートを持つ離散測度を含む.

証明. 活性点集合を  $\bar{Y}(x) := \{y \in Y \mid g(x, y) = 0\}$  とおくと, (2.4) の相補条件と  $\lambda$  の非負性より開集合  $Y \setminus \bar{Y}(x)$  上で  $\lambda = 0$  である. このとき, (2.4) の停留条件は

$$\nabla f(x) + \int_{\bar{Y}(x)} \nabla_x g(x, y) d\lambda = 0$$

となるが,  $S$  を含む最小の凸錐を  $\text{cone} S$  で表すとする, これは

$$-\nabla f(x) \in \text{cone}\{\nabla_x g(x, y) \mid y \in \bar{Y}(x)\}$$

と等価である(このことは積分論より自明としてもよいが、凸集合の強意の分離定理を用いても示せる). Carathéodory の定理(例えば Rockafellar, 1970 を参照)により、 $\mathbb{R}^n$  の部分集合  $\{\nabla_x g(x, y) \mid y \in \bar{Y}(x)\}$  の点の非負線形結合である  $-\nabla f(x)$  は高々  $n$  個の点の非負線形結合で表される. したがって、 $\lambda$  は高々  $n$  個のサポートを持つ離散測度と考えてよい.  $\square$

### 3. 測度空間における線形最適化

#### 3.1 容量問題とモーメント問題の一般化

電磁気学における静電容量問題が測度空間上の不等式制約条件つき最適化問題として定式化される (Anderson and Nash, 1987) ように、古典的なモーメント問題も測度空間上の等式制約条件つき最適化問題として様々な形で一般化され、確率論、経済学、制御理論など多くの分野に应用されている (Lasserre, 2010).

以下のようなモーメント問題の一般化を考えよう.

$$\begin{cases} \min_{\mu \in \mathcal{M}} \int_X f(x) d\mu \\ \text{subject to} \int_X \psi(x, z) d\mu = h(z) \quad \forall z \in Z \end{cases}$$

ただし、 $\mathcal{M}$  は  $M(X)$  の凸集合であり、各  $z \in Z$  に対して、関数  $\psi(\cdot, z): X \rightarrow \mathbb{R}$  は任意の  $\mu \in \mathcal{M}$  に関して可積分、また  $h(z) \in \mathbb{R}$  とする.  $Z \subset \mathbb{R}^l$  は必ずしも有限集合であるとは限らない.

応用上重要なのは  $\mathcal{M}$  が  $M \subset M(X)_+$  かつ線形不等式制約条件で規定されるとき、すなわち

$$\mathcal{M} = \left\{ \mu \in M(X)_+ \mid \int_X \varphi(x, y) d\mu \geq g(y) \quad \forall y \in Y \right\}$$

のように与えられるときである. ここで、各  $y \in Y$  に対して、関数  $\varphi(\cdot, y): X \rightarrow \mathbb{R}$  は任意の  $\mu \in M(X)_+$  に関して可積分であり、 $g(y) \in \mathbb{R}$  である.  $Y \subset \mathbb{R}^m$  も有限集合であるとは限らない. このとき、一般化モーメント問題は測度空間上の等式および不等式制約条件つき線形計画問題として以下のように書ける.

$$(3.1) \quad \begin{cases} \min_{\mu \in M(X)} \int_X f(x) d\mu \\ \text{subject to} \int_X \varphi(x, y) d\mu \geq g(y) \quad \forall y \in Y \\ \int_X \psi(x, z) d\mu = h(z) \quad \forall z \in Z \\ \mu \geq 0 \end{cases}$$

特に、 $X, Y, Z$  がコンパクトな Hausdorff 空間で、関数  $f, g, h, \varphi, \psi$  がすべて連続であるとき、双対問題は以下ようになる.

$$(3.2) \quad \begin{cases} \max_{\substack{\nu \in M(Y) \\ \xi \in M(Z)}} \int_Y g(y) d\nu + \int_Z h(z) d\xi \\ \text{subject to} \int_Y \varphi(x, y) d\nu + \int_Z \psi(x, z) d\xi \leq f(x) \quad \forall x \in X \\ \nu \geq 0 \end{cases}$$

$M(X)$  などの Banach 空間は回帰的ではないが、関数  $f, g$  等の連続性を仮定することにより回帰的な Banach 空間に準じた双対理論を展開できる(双対問題(3.2)の双対が主問題(3.1)となる

ことに注意されたい). また,  $Y, Z$  が有限集合ならば, 双対問題(3.2)は半無限線形計画問題となる.

### 3.2 双方向切除平面法

前項の一般化モーメント問題は,  $X, Y, Z$  が無限集合ならば無限計画問題となる. すなわち, 主変数  $\mu$ , 双対変数  $\nu, \xi$  とともに純然たる測度ということであるが, 離散測度の極限として最適解を求める反復法を構成することができる. 以下では, 簡単のため等式制約条件がない場合, すなわちいわゆる一般容量問題に対して, 実際に切除平面法を構成してみよう(証明等はかなり複雑になるので, 詳細は Ito, 2007 をご参照いただきたい).

連続関数  $f \in C(X), g \in C(Y)$  および  $\varphi \in C(X \times Y)$  に対して以下の問題(一般容量問題)を考える.

$$(3.3) \quad \begin{cases} \min_{\mu \in M(X)} \int_X f(x) d\mu(x) \\ \text{subject to} & \int_X \varphi(x, y) d\mu(x) \geq g(y) \quad \forall y \in Y \\ & \mu \geq 0 \end{cases}$$

この双対は次のように書ける.

$$(3.4) \quad \begin{cases} \max_{\nu \in M(Y)} \int_Y g(y) d\nu(y) \\ \text{subject to} & \int_Y \varphi(x, y) d\nu(y) \leq f(x) \quad \forall x \in X \\ & \nu \geq 0 \end{cases}$$

問題(3.3), (3.4)の可解性について, 以下のような標準的な制約想定を仮定する. このとき, 問題(3.3), (3.4)に対して最適解の存在と強双対性が保証される.

仮定 1. 次式を満たす  $\mu \in M(X)_+$  および  $\nu \in M(Y)_+$  が存在する.

$$\begin{aligned} \int_X \varphi(x, y) d\mu(x) &> g(y) \quad \forall y \in Y \\ \int_Y \varphi(x, y) d\nu(y) &< f(x) \quad \forall x \in X \end{aligned}$$

さて,  $X, Y$  を有限部分集合

$$\begin{aligned} X_k &:= \{x_1^k, x_2^k, \dots, x_{n_k}^k\} \subset X \\ Y_k &:= \{y_1^k, y_2^k, \dots, y_{m_k}^k\} \subset Y \end{aligned}$$

でそれぞれ緩和すると, 有限次元の対称形線形計画問題

$$(3.5) \quad \begin{cases} \min_{\mu \in \mathbb{R}^{n_k}} \sum_{i=1}^{n_k} f(x_i^k) \mu_i \\ \text{subject to} & \sum_{i=1}^{n_k} \varphi(x_i^k, y_j^k) \mu_i \geq g(y_j^k), \quad j = 1, 2, \dots, m_k \\ & \mu_i \geq 0, \quad i = 1, 2, \dots, n_k \end{cases}$$

$$(3.6) \quad \begin{cases} \max_{\nu \in \mathbb{R}^{m_k}} \sum_{j=1}^{m_k} g(y_j^k) \nu_j \\ \text{subject to} \quad \sum_{j=1}^{m_k} \varphi(x_i^k, y_j^k) \nu_j \leq f(x_i^k), \quad i = 1, 2, \dots, n_k \\ \nu_j \geq 0, \quad j = 1, 2, \dots, m_k \end{cases}$$

が得られ、このとき以下のような双方向切除平面法(主双対切除平面法)のアルゴリズムを構成することができる。

アルゴリズム [双方向切除平面法].

ステップ 1:  $X_1 = \{x_1^1, x_2^1, \dots, x_{n_1}^1\}$  および  $Y_1 = \{y_1^1, y_2^1, \dots, y_{m_1}^1\}$  をそれぞれ  $X, Y$  の有限部分集合 ( $n_1 = |X_1|, m_1 = |Y_1|$ ) とし,  $k=1$  とする。

ステップ 2: 緩和問題(3.5)および(3.6)の最適解  $(\mu^k, \nu^k) \in \mathbb{R}^{n_k} \times \mathbb{R}^{m_k}$  を求める。

ステップ 3:  $(\mu^k, \nu^k) \in \mathbb{R}^{n_k} \times \mathbb{R}^{m_k}$  に対して

$$\delta(\mu^k) := \min_{y \in Y} \left\{ \sum_{i=1}^{n_k} \varphi(x_i^k, y) \mu_i^k - g(y) \right\}$$

$$\gamma(\nu^k) := \max_{x \in X} \left\{ \sum_{j=1}^{m_k} \varphi(x, y_j^k) \nu_j^k - f(x) \right\}$$

を計算し、右辺の最適解をそれぞれ  $\bar{y}^k, \bar{x}^k$  とする。

ステップ 4:  $\delta(\mu^k) \geq 0$  かつ  $\gamma(\nu^k) \leq 0$  なら,  $\mu^k, \nu^k$  に対応する離散測度を問題(3.3), (3.4)の最適解として終了。

ステップ 5:  $\delta(\mu^k) < 0$  なら

$$Y_{k+1} := Y_k \cup \{\bar{y}^k\} = \{y_1^{k+1}, y_2^{k+1}, \dots, y_{m_{k+1}}^{k+1}\}, \quad m_{k+1} := m_k + 1$$

さもなければ

$$Y_{k+1} := Y_k = \{y_1^{k+1}, y_2^{k+1}, \dots, y_{m_k}^{k+1}\}, \quad m_{k+1} := m_k$$

とする。

ステップ 6:  $\gamma(\nu^k) > 0$  なら

$$X_{k+1} := X_k \cup \{\bar{x}^k\} = \{x_1^{k+1}, x_2^{k+1}, \dots, x_{n_{k+1}}^{k+1}\}, \quad n_{k+1} := n_k + 1$$

さもなければ

$$X_{k+1} := X_k = \{x_1^{k+1}, x_2^{k+1}, \dots, x_{n_k}^{k+1}\}, \quad n_{k+1} := n_k$$

とする。

ステップ 7:  $k := k + 1$  としてステップ 2 へ戻る。

ステップ 4 の  $\mu^k \in \mathbb{R}^{n_k}, \nu^k \in \mathbb{R}^{m_k}$  に対応する離散測度とは、それぞれ

$$\mu_d^k(x) = \begin{cases} \mu_i^k & \text{at } x = x_i^k, \quad i = 1, 2, \dots, n_k \\ 0 & \text{elsewhere on } X \end{cases}$$

$$\nu_d^k(y) = \begin{cases} \nu_j^k & \text{at } y = y_j^k, \quad j = 1, 2, \dots, m_k \\ 0 & \text{elsewhere on } Y \end{cases}$$

を密度関数とする  $X, Y$  上の測度を表すものとし、特に両者を区別せず  $\mu^k \in M(X), \nu^k \in M(Y)$  などとも表記することとする。

ここで仮定 1 よりも強い以下の制約想定を導入する.

仮定 2. 次式を満たす  $\mu \in \mathbb{R}^{n_1}$  および  $\nu \in \mathbb{R}^{m_1}$  が存在する (ように  $X_1, Y_1$  が選ばれている).

$$\begin{aligned} \sum_{i=1}^{n_1} \varphi(x_i^1, y) \mu_i &> g(y) \quad \forall y \in Y \\ \sum_{j=1}^{m_1} \varphi(x, y_j^1) \nu_j &< f(x) \quad \forall x \in X \end{aligned}$$

この仮定のもとでは, 緩和問題 (3.5) および (3.6) の可解性すなわち最適解の存在と強双対性が保証され (実際にはもっと弱い仮定でよい), さらに双方向切除平面法の大域的収束性が保証される.

定理 3. 仮定 2 が成り立つとし, 双方向切除平面法のアルゴリズムで生成される緩和問題 (3.5) および (3.6) の解をそれぞれ  $\mu^k, \nu^k$  とする. このとき  $M(X)$  および  $M(Y)$  で問題 (3.3), (3.4) の最適解にそれぞれ \* 弱収束する  $\{\mu^k\}, \{\nu^k\}$  の部分列が存在する.

アルゴリズムのステップ 2 において緩和問題を解く際は, 厳密に解く必要はなく,  $\Delta_k > 0$  に対して

$$\sum_{i=1}^{n_k} f(x_i^k) \mu_i^k - \sum_{j=1}^{m_k} g(y_j^k) \nu_j^k \leq \Delta_k$$

を満たす許容解  $(\mu^k, \nu^k) \in \mathbb{R}^{n_k} \times \mathbb{R}^{m_k}$  で代用できる.  $\Delta_k \rightarrow 0$  である限り大域的収束性は損なわれない.

### 3.3 確率測度の線形最適化

確率測度の凸最適化の最も簡単な例として, コンパクトな Hausdorff 空間  $X$  上の連続関数  $f \in C(X)$  に対して, 次のような線形最適化問題

$$(3.7) \quad \begin{cases} \min_{\mu \in M(X)} \int_X f(x) d\mu \\ \text{subject to} \quad \int_X d\mu = 1, \quad \mu \geq 0 \end{cases}$$

を考えることができる. 問題 (3.7) の 2 つの制約条件は  $\mu$  が  $X$  上の確率測度であることを示しており, 全体が  $f$  の  $X$  における (大域的) 最小値を求める問題となっていることを考えれば, その最適解が離散的であることはわざわざ半無限計画の理論を持ち出さなくとも納得がいくことであろう. 実際, 問題 (3.7) は半無限計画問題

$$\begin{cases} \max_{\xi \in \mathbb{R}} \xi \\ \text{subject to} \quad \xi \leq f(x) \quad \forall x \in X \end{cases}$$

と互いに双対である.

他のモーメント条件がある場合はどうか. 例えば,  $\varphi_i \in C(X)$  ( $i = 1, 2, \dots, m$ ) に対

して、不等式制約を持つ線形最適化問題

$$\begin{cases} \min_{\mu \in M(X)} \int_X f(x) d\mu \\ \text{subject to} & \int_X d\mu = 1, \quad \mu \geq 0 \\ & \int_X \varphi_i(x) d\mu \leq g_i, \quad i = 1, 2, \dots, m \end{cases}$$

は半無限計画問題

$$\begin{cases} \max_{\xi \in \mathbb{R}, \nu \in \mathbb{R}^m} \xi - \sum_{i=1}^m g_i \nu_i \\ \text{subject to} & \xi - \sum_{i=1}^m \varphi_i(x) \nu_i \leq f(x) \quad \forall x \in X \\ & \nu \geq 0 \end{cases}$$

と互いに双対であり、その最適解は高々  $m + 1$  個のサポートを持つ離散測度としてよいことがわかる。

#### 4. 測度空間における非線形凸最適化

##### 4.1 一般の凸最適化

一般の汎関数  $F$  に対して  $M(X)$  における最適化問題

$$(4.1) \quad \begin{cases} \min_{\mu \in M(X)} F(\mu) \\ \text{subject to} & \mu \in \mathcal{M} \end{cases}$$

を考える。このとき以下の命題が成り立つ。

**定理 4.**  $\mathcal{M} \subset M(X)$  が凸集合、 $F: \mathcal{M} \rightarrow \mathbb{R}$  が凸汎関数であるとする。  $\mu^* \in M(X)$  が問題 (4.1) の最適解であるための必要十分条件は、 $F$  が  $\mu^*$  において Fréchet 微分可能ならば、

$$(4.2) \quad F'(\mu^*)(\mu - \mu^*) \geq 0 \quad \forall \mu \in \mathcal{M}$$

が成り立つことである。

**証明.** 任意の  $\mu \in \mathcal{M}$  と  $\alpha \in (0, 1)$  に対して  $(1 - \alpha)\mu^* + \alpha\mu \in \mathcal{M}$  であり

$$\begin{aligned} F((1 - \alpha)\mu^* + \alpha\mu) - F(\mu^*) &= F(\mu^* + \alpha(\mu - \mu^*)) - F(\mu^*) \\ &= \alpha F'(\mu^*)(\mu - \mu^*) + o(\alpha) \end{aligned}$$

であるから、必要性は

$$0 \leq F'(\mu^*)(\mu - \mu^*) + \frac{o(\alpha)}{\alpha}$$

より、十分性は

$$\begin{aligned} F(\mu) - F(\mu^*) &= \frac{(1 - \alpha)F(\mu^*) + \alpha F(\mu) - F(\mu^*)}{\alpha} \\ &\geq \frac{F((1 - \alpha)\mu^* + \alpha\mu) - F(\mu^*)}{\alpha} \\ &= F'(\mu^*)(\mu - \mu^*) + \frac{o(\alpha)}{\alpha} \end{aligned}$$

より、それぞれ  $\alpha \rightarrow +0$  とすることにより得られる。□

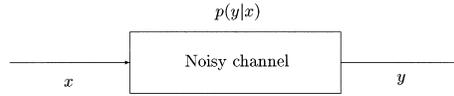


図 2. 雑音のあるスカラー入出力通信路.

不等式(4.2)は  $\mu^*$  が線形最適化問題

$$\begin{cases} \min_{\mu \in M(X)} F'(\mu^*)\mu \\ \text{subject to } \mu \in \mathcal{M} \end{cases}$$

の最適解, すなわち

$$\begin{aligned} \mu^* \in \arg \min_{\mu \in M(X)} F'(\mu^*)\mu \\ \text{subject to } \mu \in \mathcal{M} \end{aligned}$$

であることである.

#### 4.2 通信路容量問題

確率測度に関する非線形凸最適化問題の例としては通信路容量問題がある. 図 2 のような条件つき確率分布を用いて定義された雑音のあるスカラー入出力通信路を考えよう.

通信路容量とは単位時間に送ることができる情報量の上限であり, 具体的には入力信号  $x$  の確率測度  $\mu$  と通信路  $p(y|x)$  により定義される相互情報量

$$(4.3) \quad I(X, Y) := \iint p(y|x) \log \frac{p(y|x)}{p(y; \mu)} dy d\mu$$

を適当な制約条件のもとで  $\mu$  に関して最大化することによって得られる(ここで

$$p(y; \mu) := \int p(y|x) d\mu$$

とおいた).

(4.3) 式で定義される相互情報量は

$$\mathcal{M} := \{\mu \in M(X) \mid \|\mu\| = 1, \mu \geq 0\}$$

上で方向微分可能であるが, Fréchet 微分可能ではない. そこで

$$F(\mu) := \iint p(y|x) \log \frac{p(y|x)}{p(y; \mu)} dy d\mu + \iint p(y; \mu) dy d\mu - \iint p(y|x) dy d\mu$$

と修正すると,  $F$  は  $\mathcal{M}$  上で Fréchet 微分可能となる. ここで

$$\begin{aligned} \iint p(y; \mu) dy d\mu &= \int p(y; \mu) dy \int d\mu = \iint p(y|x) d\mu dy \int d\mu = \|\mu\|^2 \\ \iint p(y|x) dy d\mu &= \int d\mu = \|\mu\| \end{aligned}$$

であるから,  $\mathcal{M}$  上での値は変わらないことに注意されたい. また, 修正された相互情報量  $F$  に対しては

$$F'(\mu^*)\mu = \int \left\{ \int p(y|x) \log \frac{p(y|x)}{p(y; \mu^*)} dy + 2(\|\mu^*\| - 1) \right\} d\mu \quad \forall \mu \in M(X)$$

が成り立つ。

Smith (1971) は、特にピークパワー制約  $|x| \leq a$  が存在するとき、AWGN (Additive White Gaussian Noise) 通信路の容量を達成する入力分布が離散分布になることを示した。ピークパワー制約の存在は、入力測度がコンパクト空間上で定義されることに相当する。Smith は (4.2) を満たす  $\mu^*$  は離散的でなければならないことを複素関数論の一致の定理に基づいて証明している。このような通信路と制約条件の組合せの報告はその後も増え続けている (Chan et al., 2005; 池田, 2012)。

## 5. おわりに

本研究は日本学術振興会の科学研究費補助金 基盤研究 (C) 20540147 および挑戦的萌芽研究 24654030 の助成を受けて行われました。本稿は、日本オペレーションズ・リサーチ学会第 24 回 RAMP シンポジウムにおける講演の予稿 (伊藤, 2012) に加筆訂正を加えたものです。初稿の誤りを指摘してくださった同僚の池田思朗准教授に感謝いたします。

## 参 考 文 献

- Anderson, E. J. and Nash, P. (1987). *Linear Programming in Infinite-Dimensional Spaces: Theory and Applications*, John Wiley & Sons, Chichester.
- Berkovitz, L. D. (1961). Variational methods in problems of control and programming, *Journal of Mathematical Analysis and Applications*, **3**, 145–169.
- Chan, T. H., Hranilovic, S. and Kschischang, F. R. (2005). Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs, *IEEE Transactions on Information Theory*, **51**, 2073–2088.
- Girsanov, I. V. (1972). *Lectures on Mathematical Theory of Extremum Problems*, Lecture Notes in Economics and Mathematical Systems, **67**, Springer-Verlag, Berlin.
- 池田思朗 (2012). 通信路容量と確率測度の最適化 最適な変調方式のために, 電子通信情報学会 *Fundamentals Review*, **5**, 230–238.
- 伊藤 聡 (1998). 半無限計画の最適性と双対性, 統計数理, **46**, 345–358.
- Ito, S. (2007). Bilateral cutting plane iteration for solving general capacity problems, 「最適化：モデリングとアルゴリズム 20」, 統計数理研究所共同研究レポート, No. 203, 243–254.
- 伊藤 聡 (2012). 測度空間における凸最適化—無限次元における離散と連続, 第 24 回 RAMP シンポジウム論文集, 日本オペレーションズ・リサーチ学会常設研究部会数理計画, 59–68.
- Ito, S. and Shimizu, K. (1990). Necessary conditions for constrained optimal control problems via mathematical programming, *Numerical Functional Analysis and Optimization*, **11**, 267–281.
- Ito, S., Liu, Y. and Teo, K. L. (2000). A dual parametrization method for convex semi-infinite programming, *Annals of Operations Research*, **98**, 189–213.
- Lasserre, J. B. (2010). *Moments, Positive Polynomials and Their Applications*, Imperial College Press, London.
- Rockafellar, R. T. (1970). *Convex Analysis*, Princeton University Press, Princeton.
- Smith, J. G. (1971). The information capacity of amplitude- and variance-constrained scalar Gaussian channels, *Information and Control*, **18**, 203–219.
- 田辺広城 (1978). 『関数解析 上』, 実教出版, 東京.
- Yosida, K. (1980). *Functional Analysis*, sixth edition, Springer-Verlag, Berlin.

## Convex Optimization in Measure Spaces: Discreteness and Continuity in Infinite Dimension

Satoshi Ito

The Institute of Statistical Mathematics

Most, if not all, optimization problems in infinite dimension can be regarded as those of some measure. A measure generally has an absolutely continuous component and a discrete component with respect to Lebesgue measure. We often observe that many optimization problems in measure spaces have discrete optimal solutions. A question then arises: in what conditions does a solution to a given class of optimization problem become discrete or absolutely continuous? As an answer to such a question, this paper is concerned with convex optimization of measures defined over some compact topological spaces. This paper is a revised version of a manuscript in the Proceedings of the 24th RAMP (Research Association of Mathematical Programming) Symposium of the Operations Research Society of Japan.