凸関数とルジャンドル変換についてのノート

REIJI TOMATSU

- **おことわり**

下記ノート中の数学用語は世間に広く流布しているものではないので，ご注意下さい．

1. 単調関数の右連続化と「逆関数」

有限区間 $I=[a,b]$ を固定しておく．ここでは単調増加関数は広義のものとする．単調増加関数 φ に対して，$\varphi(x\pm 0):=\lim_{\varepsilon\to 0+}\varphi(x\pm\varepsilon)$ と書く．実数の性質からこれは必ず収束する (下限の存在).

Lemma 1.1. φ を I 上の単調増加関数とすると，I のうち高々可算個の点でジャンプが起きうる，すなわち集合 $J_\varphi:=\{t\in I\mid \varphi(t-0)<\varphi(t+0)\}$ は高々可算濃度をもつ．

Proof. 各点 $t\in J_\varphi$ に対して，$\varphi(t-0)<q(t)<\varphi(t+0)$ をみたす有理数 $q(t)$ を適当に一つ選ぶ．写像 $q:J_\varphi\to\mathbb{Q}$ が単射的であることを示す．$s,t\in J_\varphi$, $s<t$ とする (J_φ が空集合の時や 1 点のときは示すことは何もない)．すると右, 左極限の定義から $\varphi(s+0)\leq\varphi(t-0)$ である．$q(s)<\varphi(s+0)$, $\varphi(t-0)<q(t)$ だから当然 $q(s)\neq q(t)$ である．よって q は単射的であり，$\mathbb{Q}$ が可算濃度をもつから J_φ もそうである． □

特に単調増加関数は可測である．

Definition 1.2. 上の単調増加関数 φ をとる．$[a,b)$ 上の関数 $\hat{\varphi}$ を次で定義する．

$$\hat{\varphi}(t)=\varphi(t+0)\quad\text{for all } t\in[a,b).$$

$\hat{\varphi}$ を φ の**右連続化**という．

Lemma 1.3. φ を単調増加関数とするとき，その右連続化 $\hat{\varphi}$ は $[a,b)$ 上で右連続単調増加関数である．

Proof. まず単調性をチェックする．I の 2 点 $t<t'$ をとる．すると $\varphi(t+0)\leq\varphi(t')\leq\varphi(t'+0)$ であることが定義からわかる．次に右連続性をチェックする．単調性から $\inf\{\hat{\varphi}(t+\varepsilon)\mid\varepsilon>0\}=\hat{\varphi}(t)$ をチェックすることにする．これは，$\hat{\varphi}(t+\varepsilon)=\inf\{\varphi(t+\varepsilon+\delta)\mid\delta>0\}$ なので,

$$\inf\{\hat{\varphi}(t+\varepsilon)\mid\varepsilon>0\}=\inf\{\varphi(t+\varepsilon+\delta)\mid\varepsilon,\delta>0\}=\hat{\varphi}(t)$$

と変形されることから従う． □

Definition 1.4. $\varphi\colon[a,b]\to\mathbb{R}$ を単調増加関数とする．関数 $\varphi^{-1}\colon[\varphi(a),\varphi(b)]\to\mathbb{R}$ を次で定義する．

$$\varphi^{-1}(p)=\sup\{t\in[a,b]\mid\varphi(t)\leq p\}\quad\text{for all } p\in[\varphi(a),\varphi(b)].$$

これを φ の**逆関数**と呼ぶ.

φ^{-1} の性質は以下のとおり．

Lemma 1.5. φ を I 上の単調増加関数とする．このとき次が成り立つ．

(1) φ^{-1} は $[\varphi(a), \varphi(b))$ 上で右連続単調増加関数．
(2) $[\varphi^{-1}(\varphi(a)), b)$ 上で，$(\varphi^{-1})^{-1} = \hat{\varphi}$ がなりたつ．特に φ が右連続であれば，$(\varphi^{-1})^{-1} = \varphi$ となる．

Proof. (1). 単調性を確認する．2 点 $p < p'$ を区間 $[\varphi(a), \varphi(b)]$ からとる．点 $t \in I$ が $\varphi(t) \leq p$ を満たしているとすれば，$\varphi(t) \leq q'$ もいえる．よって，$\{t \mid \varphi(t) \leq p\} \subset \{t \mid \varphi(t) \leq p'\}$ がいえて，定義から $\varphi^{-1}(p) \leq \varphi^{-1}(p')$ がいえる．

次に右連続であることを示す．(φ が定数関数なら $[\varphi(a), \varphi(b))$ は空集合であるからそうでない場合を考える.) 点 $p \in [\varphi(a), \varphi(b))$ をとる．$t_p = \varphi^{-1}(p)$ と書き，$\varepsilon > 0$ を固定する．$\delta := \varphi(t_p + \varepsilon) - p \geq 0$ と δ を決める．ここで $\delta > 0$ であることに注意しておく．実際に $t_p + \varepsilon$ は (閉じているかどうかは分からない) 集合 $\{t \mid \varphi(t) \leq p\}$ の上限を ε だけ越しているから，この集合には含まれないので，特に $\varphi(t_p + \varepsilon) > p$ である．よって $\delta > 0$．さて，$q \in [\varphi(a), \varphi(b))$ が $0 \leq q - p < \delta$ を満たすならば，$|\varphi^{-1}(q) - \varphi^{-1}(p)| < \varepsilon$ であることを示そう．$p \leq q < p + \delta = \varphi(t_p + \varepsilon)$. すると定義によって，$\varphi^{-1}(q) \leq t_p + \varepsilon$. 一方 φ^{-1} は単調増加であるから，$t_p = \varphi^{-1}(p) \leq \varphi^{-1}(q)$. よって特に $|\varphi^{-1}(q) - \varphi^{-1}(p)| < \varepsilon$ である．ε は任意であるから右連続性が示された．

(2). まず，$(\varphi^{-1})^{-1} \leq \hat{\varphi}$ を示そう．定義により，

$$(\varphi^{-1})^{-1}(t) = \sup\{p \mid \varphi^{-1}(p) \leq t\}. \tag{1.1}$$

さて $\varphi^{-1}(p) \leq t$ ならば, $p \leq \hat{\varphi}(t)$ である．実際, $\varepsilon > 0$ として $t+\varepsilon$ は集合 $\{u \in I \mid \varphi(u) \leq p\}$ の上限 $\varphi^{-1}(p)$ を真に越えるから，$\varphi(t+\varepsilon) > p$ である．$\varepsilon \to 0$ として $\hat{\varphi}(t) \geq p$ がいえる．したがって (1.1) において $(\varphi^{-1})^{-1}(t) \leq \hat{\varphi}(t)$ が従う．

次に $(\varphi^{-1})^{-1} \geq \varphi$ を示す．$t \in I$，$\varepsilon > 0$ を取る．$p = \varphi(t) - \varepsilon$ とおけば，$\varphi^{-1}(p) \leq t$ であることをいう．これには，$\{u \in I \mid \varphi(u) \leq p\} \subset [a, t]$ であることをいえばよい．実際に，$p = \varphi(t) - \varepsilon$ なので，$\varphi(u) \leq p$ でかつ $u > t$ は φ の単調増加性よりありえない．よって，$\varphi^{-1}(p) \leq t$. 関数 $(\varphi^{-1})^{-1}$ の定義から，$p \leq (\varphi^{-1})^{-1}(t)$, つまり $\varphi(t) - \varepsilon = p \leq (\varphi^{-1})^{-1}(t)$. ε は任意であったから $(\varphi^{-1})^{-1} \geq \varphi$ が従う．

さて得られた関数の不等式 $\varphi \leq (\varphi^{-1})^{-1} \leq \hat{\varphi}$ のそれぞれの辺に右連続化を施す (右連続化は不等号を保存することに注意) と，φ のみ変わって $\hat{\varphi}$ となる．よって $(\varphi^{-1})^{-1} = \hat{\varphi}$. □

凸関数との関係で次は重要である．

Lemma 1.6. $\varphi\colon I \to \mathbb{R}$ を単調増加関数とする．このとき，関数 $\widetilde{\varphi}\colon I \to \mathbb{R}$ を次で定義する．

$$\widetilde{\varphi}(x) = \int_a^x \varphi(t)\, dt.$$

このとき，$\widetilde{\varphi}$ は下に凸な関数である．

Proof. 内点 $c \in I$ をとり，次式をしめす.

$$\frac{\widetilde{\varphi}(c) - \widetilde{\varphi}(a)}{c - a} \leq \frac{\widetilde{\varphi}(b) - \widetilde{\varphi}(c)}{b - c} \tag{1.2}$$

まず左辺について，

$$\begin{aligned}\widetilde{\varphi}(c)-\widetilde{\varphi}(a) &= \int_a^c \varphi(t)dt \\ &\leq \int_a^c \varphi(c)dt \\ &= (c-a)\varphi(c)\end{aligned}$$

から次が従う．

$$\frac{\widetilde{\varphi}(c)-\widetilde{\varphi}(a)}{c-a} \leq \varphi(c).$$

次に右辺について，

$$\begin{aligned}\widetilde{\varphi}(b)-\widetilde{\varphi}(c) &= \int_c^b \varphi(t)dt \\ &\geq \int_c^b \varphi(c)dt \\ &= (b-c)\varphi(c)\end{aligned}$$

から次がいえる．

$$\varphi(c) \leq \frac{\widetilde{\varphi}(b)-\widetilde{\varphi}(c)}{b-c}.$$

上の二つを合わせて，(1.2) が示された． □

2. 凸関数のルジャンドル変換

区間 $I=[a,b]$ 上の (下に) 凸な関数 f を考える．凸性は連続性を導くことを注意しておく．さて右 (左) 微分作用素 D^+, D^- を次で定義する．

$$D^+f(x)=\lim_{\varepsilon\downarrow 0}\frac{f(x+\varepsilon)-f(x)}{\varepsilon} \quad \text{for all } x\in[a,b),$$

$$D^-f(x)=\lim_{\varepsilon\downarrow 0}\frac{f(x)-f(x-\varepsilon)}{\varepsilon} \quad \text{for all } x\in(a,b].$$

これらが定義できることは，lim 内が ε に対して二つとも単調に振舞うことからわかる．凸性から導かれる次の不等式は基本的である．

$$D^-f(x)\leq D^+f(x)\leq D^-f(y)\leq D^+f(y) \quad \text{for all } a<x<y<b.$$

$a\leq x<y\leq b$ についても真ん中の不等式 $D^+f(x)\leq D^-f(y)$ は成り立つことにも注意しよう．これ以降，区間の端での微分は有限値をもつものと仮定する:

$$-\infty<D^+f(a), \quad D^-f(b)<\infty.$$

Lemma 2.1. I 上の凸関数 f にたいして，次が成り立つ．

(1) $D^+f:[a,b)\to\mathbb{R}$ は右連続単調増加関数．
(2) $D^-f:(a,b]\to\mathbb{R}$ は左連続単調増加関数．

Proof. (1). $x \in [a,b), \varepsilon > 0$をとる．それにたいして，(極限の定義より)$\delta > 0$を$x+\delta < b$かつ

$$\frac{f(x+\delta)-f(x)}{\delta} < D^+f(x)+\varepsilon$$

となるようにとってこれる．関数$x \mapsto f(x+\delta)-f(x)$は連続(凸性が連続性を導くことに注意)だから，次を満たすような$\delta' > 0$が存在する．

$$\frac{f(x+\delta+\delta')-f(x+\delta')}{\delta} < D^+f(x)+\varepsilon.$$

定義により次が成り立つ．

$$D^+f(x+\delta') < D^+f(x)+\varepsilon.$$

よって，

$$D^+f(x) \leq \lim_{\mu\downarrow 0} D^+f(x+\mu) = \inf_{\mu>0} D^+f(x+\mu) \leq D^+f(x+\delta') < D^+f(x)+\varepsilon.$$

不等式

$$D^+f(x) \leq \lim_{\mu\downarrow 0} D^+f(x+\mu) < D^+f(x)+\varepsilon$$

についてはεは任意に動かせるので，$D^+f(x) = \lim_{\mu\downarrow 0} D^+f(x+\mu)$, つまり右連続である．

(2). (1)と同様. □

前にLemma 1.6で，単調増加関数を積分すると凸関数になることを示したが，任意の凸関数はそのように記述できることをみよう．

Lemma 2.2. fをI上の凸関数とする．このときfは次のように積分表示される．

$$f(x) = f(a) + \int_a^x D^+f(t)dt.$$

ここでD^+fをD^-fに変更しても等式はなりたつ．

Proof. D^+fの方の等式を示す．D^-fの方も同様に証明できる．右微分の定義から$a < t < x$をみたすtについて，次がいえる．

$$D^+f(t) = \lim_{n\to\infty} n(f(t+1/n)-f(t)).$$

この lim 内の数値はnに関わらず上から$D^-f(b) < \infty$で押さえられるので，ルベーグの収束定理によって極限と積分を入れかえた式が成り立つ:

$$\int_a^x D^+f(t)dt = \lim_{n\to\infty}\int_a^x n(f(t+1/n)-f(t))dt.$$

右側の式は以下のように計算できるので，この補題にある式がなりたつ．

$$\begin{aligned}\lim_{n\to\infty}\int_a^x n(f(t+1/n)-f(t))dt &= \lim_{n\to\infty} n\int_{a+1/n}^{x+1/n} f(t)dt - n\int_a^x f(t)dt \\ &= \lim_{n\to\infty} n\int_x^{x+1/n} f(t)dt - n\int_a^{a+1/n} f(t)dt \\ &= f(x)-f(a),\end{aligned}$$

ここで最後の等式は，fの連続性から従う． □

Lemma 2.3. I上の凸関数fについて次が成り立つ．

(1)

$$D^+f(x-0) = D^-f(x) \quad \textit{for all } x \in (a,b)$$

(2)

$$D^-f(x+0) = D^+f(x) \quad \textit{for all } x \in (a,b)$$

Proof. (1). 凸関数 f は単調増加関数 φ を用いて

$$f(x) = f(a) + \int_a^x \varphi(t)dt$$

と積分表示される (たとえば $\varphi = D^+f$ あるいは $\varphi = D^-f$). すると次がいえる.

$$D^+f(x) = \varphi(x+0), \quad D^-f(x) = \varphi(x-0) \tag{2.1}$$

右側は左側と同じように済むので左側だけ示す. $\varepsilon > 0$ とする. $\varphi(t_0) \le \varphi(x+0) + \varepsilon$ となる $x < t_0$ をとってくる. するともちろん $x < t < t_1 \le t_0$ では $\varphi(t) \le \varphi(x+0) + \varepsilon$. よって

$$\begin{aligned} \frac{f(t_1) - f(x)}{t_1 - x} &= (t_1 - x)^{-1} \int_x^{t_1} \varphi(t)dt \\ &\le (t_1 - x)^{-1} \int_x^{t_1} (\varphi(x+0) + \varepsilon)dt \\ &= \varphi(x+0) + \varepsilon \end{aligned} \tag{2.2}$$

となり, これから

$$\frac{f(t_1) - f(x)}{t_1 - x} \le \varphi(x+0) + \varepsilon.$$

(2.2) に $\varphi(t) \ge \varphi(x+0)$ を使うと, 次の下からの評価を得る

$$\varphi(x+0) \le \frac{f(t_1) - f(x)}{t_1 - x}.$$

$x < t_1 \le t_0$ で $t_1 \to x$ の極限をとると,

$$\varphi(x+0) \le D^+f(x) \le \varphi(x+0) + \varepsilon.$$

これから $x \in (a,b)$ に対して, $\varphi(x+0) = D^+f(x)$ が分かった. 同様に $\varphi(x-0) = D^-f(x)$ もいえる.

さて, (2.1) の左側の式で $\varphi = D^-f$ としてみると, $D^+f(x) = D^-f(x+0)$ がいえる. また (2.1) の右側の式で $\varphi = D^+f$ としてみると, $D^-f(x) = D^+f(x-0)$ がいえる.. □

さてルジャンドル変換の定義をしよう. そのために, まず

$$\alpha = \inf_{t \in (a,b)} D^-f(t), \quad \beta = \sup_{t \in (a,b)} D^+f(t).$$

とおく. 我々はこれらが実数値であるものについて話を進めている. 前補題によりこれらは次のようにも書ける.

$$\alpha := D^+f(a), \quad \beta =: D^-f(b)$$

Definition 2.4. 以上の設定のもとで, f のルジャンドル変換を次で定める.

$$g(p) := p(D^+f)^{-1}(p) - f((D^+f)^{-1}(p)) \quad \text{for all } p \in (\alpha, \beta).$$

$g(p)$ の積分表示を与えよう．$\alpha = D^+f(a) < p < D^-f(b) = \beta$ と仮定しよう．$c := (D^+f)^{-1}(p)$ とおき，長方形集合 $R = [a, c] \times [\alpha, p]$ を考える．今自明なことだが，次がなりたつ．

$$R = \{(x, y) \in R \mid D^+f(x) \geq y\} \cup \{(x, y) \in R \mid D^+f(x) < y\}.$$

μ を $\mathbb{R}^2$ のルベーグ測度とする. 集合 $A = \{(x, y) \in R \mid D^+f(x) < y\}$，$B = \{(x, y) \in R \mid x \leq (D^+f)^{-1}(y)\}$ をくらべて，$A \subset B$ であり両者の測度が変わらないことを示す．$A \subset B$ は単調増加関数の逆関数の定義から明らか. $(x, y) \in B$ なら $x \leq (D^+f)^{-1}(y)$ であるので，$D^-f(x) = D^+f(x-0) \leq y$. よって $B \setminus A$ は可測集合 $\{(x, y) \in R \mid D^-f(x) \leq y \leq D^+f(x)\}$ に含まれる．これが測度零であることは，Fubini の定理を使って次のようにして分かる (χ_E は集合 E の特性関数，つまり χ_E の値域は $\{0, 1\}$ で，$\chi_E(x) = 1$ となるのは $x \in E$ に対してのみ，という関数).

$$\begin{aligned}
\mu(\{(x, y) \in R \mid D^-f(x) \leq y \leq D^+f(x)\}) &= \int_a^c dx \left(\int_\alpha^p dy \chi_{\{(x,y) \in R \mid D^-f(x) \leq y \leq D^+f(x)\}} \right) \\
&= \int_a^c dx \, (D^+f(x) - D^-f(x)) \\
&= f(c) - f(a) - (f(c) - f(a)) = 0,
\end{aligned}$$

ここで最後の等式は, Lemma 2.2 から従う．よって $\mu(A) = \mu(B)$.

一方

$$\mu(B) = \int_\alpha^p ((D^+f)^{-1}(y) - a) \, dy$$

より，

$$\begin{aligned}
\mu(R) &= \mu(\{(x, y) \in R \mid D^+f(x) \geq y\}) + \mu(\{(x, y) \in R \mid D^+f(x) < y\}) \\
&= \mu(\{(x, y) \in R \mid D^+f(x) \geq y\}) + \mu(A) \\
&= \mu(\{(x, y) \in R \mid D^+f(x) \geq y\}) + \mu(B) \\
&= \int_a^c (D^+f(x) - \alpha) \, dx + \int_\alpha^p ((D^+f)^{-1}(y) - a) \, dy \\
&= f(c) - f(a) - \alpha(c - a) + \int_\alpha^p ((D^+f)^{-1}(y) - a) \, dy \\
&= f(c) - f(a) - \alpha(c - a) - a(p - \alpha) + \int_\alpha^p (D^+f)^{-1}(y) \, dy.
\end{aligned}$$

一方 $\mu(R) = (c - a)(p - \alpha)$, $g(p) = cp - f(c)$ なので，次が得られた．

Lemma 2.5. f は $[a, b]$ 上の凸関数で，$\alpha := \inf\{D^-f(t) \mid t \in I\}$, $\beta := \sup\{D^+f(t) \mid \in I\}$ は有限値とする．g を f のルジャンドル変換とすると，$D^+f(a) < p < D^+f(b)$ となる p に対して，次が成り立つ．

$$g(p) = a\alpha - f(a) + \int_\alpha^p (D^+f)^{-1}(y) \, dy.$$

特に, g は凸関数であり

$$D^-g(p) = (D^+f)^{-1}(p - 0), \quad D^+g(p) = (D^+f)^{-1}(p + 0) = D^+f(p).$$

微係数のジャンプについては次がいえる．

Lemma 2.6. 上の補題と同じ状況で，次がなりたつ．

$$\{t \in (a,b) \mid D^- g(p) \le t \le D^+ g(p)\} = \{t \in (a,b) \mid D^+ f(t-0) \le p \le D^+ f(t)\}.$$

Proof. $D^- g(p) \le t \le D^+ g(p)$ とする．これは

$$(D^+ f)^{-1}(p-0) \le t \le (D^+ f)^{-1}(p)$$

を意味する．左側の不等号は任意の $\varepsilon, \delta > 0$ について，$(D^+ f)^{-1}(p-\varepsilon) < t+\delta$ を導く．従って，$p - \varepsilon < D^+ f(t+\delta)$．すなわち，$p \le D^+ f(t+0) = D^+ f(t)$ がいえる．次に右側の不等式から $D^+ f(t-0) \le p$ がいえる．よって包含関係 $\subset$ がなりたつ．逆をたどれば $\supset$ もいえる． □

$D^- g(p) < D^+ g(p)$ の場合，$t_1 < t < t_2$ を満たす t_1, t, t_2 を $[D^- g(p), D^+ g(p)]$ からとると，$p \le D^+ f(t_1) \le D^+ f(t-0) \le p$ なので

$$D^+ f(t_1) = p = D^+ f(t-0).$$

同様に

$$D^+ f(t) = p = D^+ f(t_2 - 0).$$

よって，$D^+ f(t) = p = D^+ f(t-0) = D^- f(t)$ となる．特に f は $(D^- g(p), D^+ g(p))$ において, 微分可能で微係数は p をとる．

Department of Mathematical Sciences, University of Tokyo, Komaba, 153-8914

E-mail address: tomatsu@ms.u-tokyo.ac.jp