

# 第1章 凸集合と凸関数の基本

## 1 アフィン接続と凸性

$M$  を多様体、 $\nabla$  を  $M$  上のアフィン接続とする。

**定義 1.1** (平坦アフィン接続).  $M$  の開部分集合  $O \subset^{\text{open}} M$  上の座標であって、それに関する  $\nabla$  の接続係数がすべて 0 となるものを、 $O$  上の  **$\nabla$ -アフィン座標 ( $\nabla$ -affine coordinates)** という。

各  $p \in M$  に対し、 $p$  のまわりの  $\nabla$ -アフィン座標が存在するとき、 $\nabla$  は  $M$  上**平坦 (flat)** であるという。

**命題-定義 1.2** ( $U$  上の標準的な平坦アフィン接続). **[TODO] 書き方を修正**  $U \subset^{\text{open}} M$  とする。 $U$  上のアフィン接続  $D: \Gamma(TU) \rightarrow \Gamma(T^{\vee}U \otimes TU)$  を、次の規則で well-defined に定めることができる:

- 各  $X \in \Gamma(TU)$  に対し、 $W$  の基底が定める  $U$  上の座標  $x^i$  ( $i = 1, \dots, m$ ) をひとつ選び、

$$DX := dX^i \otimes \frac{\partial}{\partial x^i} \in \Gamma(T^{\vee}U \otimes TU) \quad (1.1)$$

と定める。ただし、 $X$  の成分表示を  $X = X^i \frac{\partial}{\partial x^i}$  とおいた。

さらに、この  $D$  は  $U$  上のアフィン接続として平坦である。 $D$  を  $U$  上の**標準的な平坦アフィン接続 (standard flat affine connection)** という。

**証明** 写像として well-defined であることを一旦認め、先に  $\mathbb{R}$ -線型性、Leibniz 則、平坦性を確かめる。 $D$  の  $\mathbb{R}$ -線型性と Leibniz 則は、外微分  $d$  の  $\mathbb{R}$ -線型性と Leibniz 則から従う。平坦性は、式 (1.1) で用いた座標  $x^i$  が  $D$ -アフィン座標となることから従う。最後に、 $D$  が写像として well-defined であることを示す。 $y^{\alpha}$  ( $\alpha = 1, \dots, m$ ) を  $W$  の基底が定める  $U$  上の座標とすると、

$$dX^i \otimes \frac{\partial}{\partial x^i} = d \left( X^{\alpha} \frac{\partial x^i}{\partial y^{\alpha}} \right) \otimes \frac{\partial y^{\alpha}}{\partial x^i} \frac{\partial}{\partial y^{\alpha}} \quad (1.2)$$

$$= \left( \frac{\partial x^i}{\partial y^{\alpha}} dX^{\alpha} + \underbrace{X^{\alpha} d \left( \frac{\partial x^i}{\partial y^{\alpha}} \right)}_{=0} \right) \otimes \frac{\partial y^{\alpha}}{\partial x^i} \frac{\partial}{\partial y^{\alpha}} \quad (1.3)$$

$$= dX^{\alpha} \otimes \frac{\partial}{\partial y^{\alpha}} \quad (1.4)$$

となる。ただし「 $= 0$ 」の部分は  $x^i$  と  $y^{\alpha}$  の間の座標変換がアフィン変換となることを用いた。これで well-defined 性も示された。  $\square$

**定義 1.3** ( $\nabla$ -凸集合). 部分集合  $S \subset M$  が  **$\nabla$ -凸 ( $\nabla$ -convex)** であるとは、任意の  $p, q \in S$  に対し、 $p$  から  $q$  への  $S$  内の  $\nabla$ -測地線がただひとつ存在することをいう。

**定義 1.4** ( $\nabla$ -凸関数).  $U \overset{\text{open}}{\subset} M$  を  $\nabla$ -凸開集合とする。関数  $f \in C^\infty(U)$  が  **$\nabla$ -凸 ( $\nabla$ -convex)** であるとは、 $U$  内の任意の  $\nabla$ -測地線  $\gamma: [0, 1] \rightarrow U$  に対し、 $f \circ \gamma: [0, 1] \rightarrow \mathbb{R}$  が凸関数であることをいう。

## 2 Hessian

$W$  を  $m$  次元  $\mathbb{R}$ -ベクトル空間 ( $m \in \mathbb{Z}_{\geq 0}$ )、 $U \overset{\text{open}}{\subset} W$  を開部分集合、 $D$  を  $U$  上の標準的な平坦アファイン接続とする。

**定義 2.1** (Hessian).  $C^\infty$  関数  $f: U \rightarrow \mathbb{R}$  に対し、 $f$  の **Hessian** を

$$\text{Hess } f := Ddf \in \Gamma(T^\vee U \otimes T^\vee U) \quad (2.1)$$

と定義する。

$D$ -アファイン座標を用いると、Hessian の成分表示は簡単な形になる。

**命題 2.2** (Hessian の成分表示).  $x^i$  ( $i = 1, \dots, m$ ) を  $U$  上の  $D$ -アファイン座標とする。このとき、座標  $x^i$  に関する  $\text{Hess } f$  の成分表示は

$$\text{Hess } f = \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i \otimes dx^j \quad (2.2)$$

となる。とくに  $f$  の  $C^\infty$  性より  $\text{Hess } f$  は対称テンソルである。

**証明**  $(\text{Hess } f)(\partial_i, \partial_j) = \langle D_{\partial_i} df, \partial_j \rangle = \partial_i \langle df, \partial_j \rangle - \underbrace{\langle df, D_{\partial_i} \partial_j \rangle}_{=0} = \partial_i (\partial_j f) = \frac{\partial^2 f}{\partial x^i \partial x^j}$  より従う。  $\square$

## 3 Legendre 変換

**定義 3.1** (Legendre 変換).  $U \subset W$  を開集合、 $f: U \rightarrow \mathbb{R}$  を  $C^\infty$  関数であって  $\nabla f: U \rightarrow W^\vee$  が単射であるものとする。関数

$$f^\vee: U' \rightarrow \mathbb{R}, \quad y \mapsto \langle (\nabla f)^{-1}(y), y \rangle - f((\nabla f)^{-1}(y)) \quad \text{where } U' := \nabla f(U) \quad (3.1)$$

を  $f$  の **Legendre 変換 (Legendre transform)** という。

**例 3.2** (Legendre 変換の例). 具体的な指数型分布族に対し、対数分配関数の Legendre 変換を計算してみる。

- Bernoulli 分布族 (i.e. 2 元集合上の full support な確率分布の族): 対数分配関数は  $\psi: \mathbb{R} \rightarrow \mathbb{R}$ ,  $\theta \mapsto \log(1 + \exp \theta)$  であった。よって  $\nabla \psi(\theta) = \frac{\exp \theta}{1 + \exp \theta}$  であり、 $(\nabla \psi)^{-1}(\eta) = \log \eta - \log(1 - \eta)$  である。したがって  $\psi^\vee(\eta) = \eta \log \eta + (1 - \eta) \log(1 - \eta)$  である。

- 正規分布族: 対数分配関数は  $\psi: \mathbb{R} \times \mathbb{R}_{<0} \rightarrow \mathbb{R}$ ,  $\theta \mapsto -\frac{(\theta^1)^2}{4\theta^2} - \frac{1}{2} \log(-\theta^2) + \frac{1}{2} \log \pi$  であった。よって  $\nabla \psi(\theta) = \left( -\frac{\theta^1}{2\theta^2}, \frac{(\theta^1)^2}{4(\theta^2)^2} - \frac{1}{2\theta^2} \right)$  であり、 $(\nabla \psi)^{-1}(\eta) = \frac{1}{\eta_2 - (\eta_1)^2} \begin{pmatrix} \eta_1 \\ -1/2 \end{pmatrix}$  である。したがって

$$\psi^\vee(\eta) = -\frac{1}{2} \left( 1 + \log 2\pi + \log(\eta_2 - (\eta_1)^2) \right) \text{ である。}$$

本稿では、とくに次の状況を考えることになる。

**命題 3.3.** [TODO] 単射の証明などは補題に切り出す  $U \subset W$  を凸開集合、 $f: U \rightarrow \mathbb{R}$  を  $C^\infty$  関数であって  $\text{Hess } f$  が  $U$  上各点で (対称であることも含む意味で) 正定値であるものとする。このとき、次が成り立つ:

- (1)  $\nabla f$  は局所微分同相である。とくに  $U' := \nabla f(U)$  は  $W^\vee$  の開集合である。
- (2)  $\nabla f: U \rightarrow U'$  は微分同相である。とくに  $\nabla f$  は単射である。

したがって  $f^\vee$  が定義でき、 $f^\vee$  は次をみたす:

- (3)  $f^\vee: U' \rightarrow \mathbb{R}$  は  $C^\infty$  関数である。
- (4)  $\nabla f^\vee = (\nabla f)^{-1}$  が成り立つ。とくに  $\nabla f^\vee$  は単射である。
- (5) 各  $y \in U'$  に対し  $(\text{Hess } f^\vee)_y = ((\text{Hess } f)_x)^{-1}$  が成り立つ (ただし  $x := (\nabla f)^{-1}(y)$ )。とくに  $(\text{Hess } f^\vee)_y$  は正定値である。

**証明** (1) 命題の仮定より  $\text{Hess } f$  は  $U$  上各点で正定値だから、 $\nabla f$  の微分は各点で線型同型である。したがって  $\nabla f$  は局所微分同相であり、とくに開写像である。よって  $U' = \nabla f(U)$  は  $W^\vee$  の開集合である。

(2)  $u, \tilde{u} \in U$ ,  $u \neq \tilde{u}$  を固定し、 $[0, 1]$  を含む  $\mathbb{R}$  の開区間  $I$  であって、すべての  $t \in I$  に対し  $(1-t)u + t\tilde{u}$  が  $U$  に属するようなものをひとつ選ぶ ( $U$  は  $W$  の凸開集合だからこれは可能)。さらに  $\varphi: I \rightarrow U$ ,  $t \mapsto f((1-t)u + t\tilde{u})$  と定めると、平均値定理より、ある  $\tau \in (0, 1)$  が存在して

$$\langle \nabla f(\tilde{u}) - \nabla f(u), \tilde{u} - u \rangle = \varphi'(1) - \varphi'(0) \quad (3.2)$$

$$= \varphi''(\tau) \quad (\text{平均値定理}) \quad (3.3)$$

$$= \langle (\text{Hess } f)_{(1-\tau)u + \tau\tilde{u}}, (\tilde{u} - u)^2 \rangle \quad (3.4)$$

$$> 0 \quad (\text{Hess } f \text{ は正定値}) \quad (3.5)$$

が成り立つ。よって  $\nabla f(\tilde{u}) \neq \nabla f(u)$  である。したがって  $\nabla f$  は単射である。このことと (1) より  $\nabla f: U \rightarrow U'$  は微分同相である。

(3)  $\nabla f: U \rightarrow U'$  が微分同相ゆえに  $(\nabla f)^{-1}: U' \rightarrow U$  は  $C^\infty$  だから、 $f^\vee$  は  $C^\infty$  関数である。

(4)  $f^\vee$  の定義式を  $\nabla$  で微分すると、すべての  $y \in U'$  に対し

$$(\nabla f^\vee)(y) = (\nabla f)^{-1}(y) + \langle y, \nabla(\nabla f)^{-1}(y) \rangle - \langle (\nabla f)((\nabla f)^{-1}(y)), \nabla(\nabla f)^{-1}(y) \rangle = (\nabla f)^{-1}(y) \quad (3.6)$$

が成り立つ。よって  $(\nabla f)^{-1} = \nabla f^\vee$  である。

(5) (4) より

$$(\text{Hess } f^\vee)_y = d(\nabla f^\vee)_y \quad (3.7)$$

$$= d((\nabla f)^{-1})_y \quad (3.8)$$

$$= (d(\nabla f)_x)^{-1} \quad (3.9)$$

$$= ((\text{Hess } f)_x)^{-1} \quad (3.10)$$

となる。

□

**系 3.4** (Legendre 変換の対合性).  $f^{\vee\vee} = f$ .

**証明** Legendre 変換の定義より、すべての  $x \in U$  に対し

$$f^{\vee\vee}(x) = \langle x, (\nabla f^{\vee})^{-1}(x) \rangle - f^{\vee}((\nabla f^{\vee})^{-1}(x)) \quad (3.11)$$

$$= \langle x, \nabla f(x) \rangle - f^{\vee}(\nabla f(x)) \quad (\nabla f^{\vee} = (\nabla f)^{-1}) \quad (3.12)$$

$$= \langle x, \nabla f(x) \rangle - \left( \langle \nabla f(x), (\nabla f)^{-1}(\nabla f(x)) \rangle - f((\nabla f)^{-1}(\nabla f(x))) \right) \quad (3.13)$$

$$= f(x) \quad (3.14)$$

が成り立つ。よって  $f^{\vee\vee} = f$  である。  $\square$

## 4 Fourier-Laplace 変換

[TODO] ちゃんと書く。cf. [?] ]

**定義 4.1** (Fourier-Laplace 変換).  $V$  を有限次元  $\mathbb{R}$ -ベクトル空間、 $\mu$  を  $V$  上の測度とする。

$$L_{\mu}(\theta) := \int_{v \in V} e^{\langle \theta, v \rangle} d\mu(v) \quad (\theta \in V^{\vee} \otimes \mathbb{C}) \quad (4.1)$$

と定め、 $L_{\mu}$  を **Fourier-Laplace 変換 (Fourier-Laplace transform)** という。