正規分布からなる定曲率統計多様体

mico

2025 年 3 月 23 日

1　はじめに

この PDF では定曲率統計多様体の典型的な例である, (1 次元の) 正規分布の空間からなる統計多様体について具体的な計算を通じて説明を行う.

2　準備

2.1　正規分布の特性関数

定義 1: 正規分布

平均 μ , 標準偏差 σ の正規分布 $p(x;\mu,\sigma)$ は

$$p(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

で与えられる.

以下, 確率変数 Y が平均 0 , 標準偏差 σ の正規分布に従う, すなわち $Y \sim N(\mu,\sigma^2)$ であるとして議論を行う.

命題 1: 正規分布の特性関数

等式

$$E[\exp(itY)] = \exp\left(-\frac{\sigma^2 t^2}{2}\right) \tag{1}$$

が成立する.

Proof.

$$
\begin{aligned}
E[\exp(itY)] &= \int_{-\infty}^{\infty} \exp(itx) p(x; 0, \sigma^2) dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2} + itx\right) dx.
\end{aligned}
$$

ここで,

$$
\begin{aligned}
-\frac{x^2}{2\sigma^2} + itx &= -\frac{1}{2\sigma^2}(x^2 - 2i\sigma^2 tx) \\
&= -\frac{1}{2\sigma^2}\left\{(x - i\sigma^2 t)^2 + \sigma^4 t^2\right\} \\
&= -\frac{(x - i\sigma t)^2}{2\sigma^2} - \frac{\sigma^2 t^2}{2}
\end{aligned}
$$

より,

$$
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2} + itx\right) dx = \exp\left(-\frac{\sigma^2 t^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x - i\sigma t)^2}{2\sigma^2}\right) dx.
$$

ここで, 複素 Gauss 積分により,

$$
\begin{aligned}
\exp\left(-\frac{\sigma^2 t^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x - i\sigma t)^2}{2\sigma^2}\right) dx &= \exp\left(-\frac{\sigma^2 t^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \sqrt{2\pi\sigma^2} \\
&= \exp\left(-\frac{\sigma^2 t^2}{2}\right).
\end{aligned}
$$

したがって,

$$
E[\exp(itY)] = \exp\left(-\frac{\sigma^2 t^2}{2}\right)
$$

□

この特性関数を用いて, $E[Y^k]$ を求めてゆく. (1) の両辺を t で微分すると,

$$
E[iY \exp(itY)] = -\sigma^2 t \exp\left(-\frac{\sigma^2 t^2}{2}\right). \tag{2}
$$

さらに t で微分すると,

$$
E[-Y^2 \exp(itY)] = \sigma^2(\sigma^2 t^2 - 1) \exp\left(-\frac{\sigma^2 t^2}{2}\right). \tag{3}
$$

繰り返していき,

$$
E[-iY^3 \exp(itY)] = \sigma^4 t(3 - \sigma^2 t^2) \exp\left(-\frac{\sigma^2 t^2}{2}\right), \tag{4}
$$

$$
E[Y^4 \exp(itY)] = \sigma^4(\sigma^4 t^4 - 6\sigma^2 t^2 + 3) \exp\left(-\frac{\sigma^2 t^2}{2}\right). \tag{5}
$$

$$E[iY^5] = \sigma^6 t(10\sigma^2 t^2 - 15 - \sigma^4 t^4) \exp\left(-\frac{\sigma^2 t^2}{2}\right). \tag{6}$$

$$E[-Y^6] = \sigma^6(45\sigma^2 t^2 - 15 - 15\sigma^4 t^4 + \sigma^6 t^6) \exp\left(-\frac{\sigma^2 t^2}{2}\right). \tag{7}$$

(2) の両辺に $t = 0$ を代入すると,

$$E[iY] = 0$$

となり, 整理して

$$E[Y] = 0$$

を得る. 同様に (3), (4), (5), (6), (7) の両辺に $t = 0$ を代入して整理すると,

$$E[Y^2] = \sigma^2$$

$$E[Y^3] = 0$$

$$E[Y^4] = 3\sigma^4$$

$$E[Y^5] = 0$$

$$E[Y^6] = 15\sigma^6$$

を得る.

さて, 確率変数 X が $X \sim N(\mu, \sigma^2)$ であるとする. このとき $Y = X - \mu$ が成立するので,

$$E[X - \mu] = 0 \tag{8}$$

$$E[(X - \mu)^2] = \sigma^2 \tag{9}$$

$$E[(X - \mu)^3] = 0 \tag{10}$$

$$E[(X - \mu)^4] = 3\sigma^4 \tag{11}$$

$$E[(X - \mu)^5] = 0 \tag{12}$$

$$E[(X - \mu)^6] = 15\sigma^6 \tag{13}$$

である.

3 α 接続の計算

3.1 Fisher 計量

$$M := \{p(x; \mu, \sigma) | \mu \in \mathbb{R}, \sigma > 0\}$$

とする. M は明らかに 2 次元多様体である. M の Fisher 計量を求める.

Fisher 計量を求める準備として, あらかじめ $\dfrac{\partial}{\partial \mu} \log p(x; \mu, \sigma)$, $\dfrac{\partial}{\partial \sigma} \log p(x; \mu, \sigma)$ を計算しておく.

以降は $\dfrac{\partial}{\partial\mu}$ を ∂_μ , $\dfrac{\partial}{\partial\sigma}$ を ∂_σ と省略して書くこととする.

$$\partial_\mu \log p(x;\mu,\sigma) = \frac{x-\mu}{\sigma^2},$$
$$\partial_\sigma \log p(x;\mu,\sigma) = -\frac{1}{\sigma} + \frac{(x-\mu)^2}{\sigma^3}$$

である.

実際に Fisher 計量 g を求めてゆく.

$$\begin{aligned} g\left(\partial_\mu, \partial_\mu\right) &= \int_{-\infty}^{\infty} p(x;\mu,\sigma)\partial_\mu \log p(x;\mu,\sigma)\partial_\mu \log p(x;\mu,\sigma)dx \\ &= \frac{1}{\sigma^4}\int_{-\infty}^{\infty}(x-\mu)^2 p(x;\mu,\sigma)dx \\ &= \frac{1}{\sigma^4}\sigma^2 = \frac{1}{\sigma^2}. \end{aligned}$$

ここで (9) を用いた.

$$\begin{aligned} g\left(\partial_\mu, \partial_\sigma\right) &= \int_{-\infty}^{\infty} p(x;\mu,\sigma)\partial_\mu \log p(x;\mu,\sigma)\partial_\sigma \log p(x;\mu,\sigma)dx \\ &= -\frac{1}{\sigma^3}\int_{-\infty}^{\infty}(x-\mu)p(x;\mu,\sigma)dx - \frac{1}{\sigma^4}\int_{-\infty}^{\infty}(x-\mu)^3 p(x;\mu,\sigma)dx \\ &= 0. \end{aligned}$$

ここで (8), (10) を用いた.

$$\begin{aligned} g\left(\partial_\sigma, \partial_\sigma\right) &= \int_{-\infty}^{\infty} p(x;\mu,\sigma)\partial_\sigma \log p(x;\mu,\sigma)\partial_\sigma \log p(x;\mu,\sigma)dx \\ &= \frac{1}{\sigma^2}\int_{-\infty}^{\infty} p(x;\mu,\sigma)dx - \frac{2}{\sigma^4}\int_{-\infty}^{\infty}(x-\mu)^2 p(x;\mu,\sigma)dx \\ &\quad + \frac{1}{\sigma^6}\int_{-\infty}^{\infty}(x-\mu)^4 p(x;\mu,\sigma)dx \\ &= \frac{1}{\sigma^2} - \frac{2}{\sigma^4}\sigma^2 + \frac{1}{\sigma^6}3\sigma^4 = \frac{2}{\sigma^2}. \end{aligned}$$

ここで (9), (11) を用いた.

以上で Fisher 情報行列の計算は終わった.

3.2 Levi-Civita 接続

Koszul の公式を用いて Levi-Civita 接続 ∇^g の計算を行う.

$$2g(\nabla^g_{\partial_\mu}\partial_\mu, \partial_\mu) = 0,$$

$$2g(\nabla^g_{\partial_\mu}\partial_\mu, \partial_\sigma) = -\partial_\sigma g(\partial_\mu, \partial_\mu) = \frac{2}{\sigma^3}$$

なので,

$$\nabla^g_{\partial_\mu}\partial_\mu = \frac{1}{2\sigma}\partial_\sigma.$$

$$2g(\nabla^g_{\partial_\mu}\partial_\sigma, \partial_\mu) = \partial_\sigma g(\partial_\mu, \partial_\mu) = -\frac{2}{\sigma^3},$$
$$2g(\nabla^g_{\partial_\mu}\partial_\sigma, \partial_\sigma) = 0$$

なので,

$$\nabla^g_{\partial_\mu}\partial_\sigma = -\frac{1}{\sigma}\partial_\mu.$$

∇^g は捩れを持たないので,

$$\nabla^g_{\partial_\sigma}\partial_\mu = \nabla^g_{\partial_\mu}\partial_\sigma = -\frac{1}{\sigma}\partial_\mu.$$

$$2g(\nabla^g_{\partial_\sigma}\partial_\sigma, \partial_\mu) = 0,$$
$$2g(\nabla^g_{\partial_\mu}\partial_\sigma, \partial_\sigma) = \partial_\sigma g(\partial_\sigma, \partial_\sigma) = -\frac{4}{\sigma^3}$$

なので,

$$\nabla^g_{\partial_\sigma}\partial_\sigma = -\frac{1}{\sigma}\partial_\sigma.$$

以上で Levi-Civita 接続の計算は終わった.

3.3 対称テンソル場 S

3 次の対称テンソル場

$$S(X, Y, Z) := E[(X \log p(x;\mu,\sigma))(Y \log p(x;\mu,\sigma))(Z \log p(x;\mu,\sigma))]$$

を計算する.

$$\begin{aligned} S(\partial_\mu, \partial_\mu, \partial_\mu) &= \int_{-\infty}^{\infty} (\partial_\mu \log p(x;\mu,\sigma))^3 p(x;\mu,\sigma)dx \\ &= \frac{1}{\sigma^6}\int_{-\infty}^{\infty} (x-\mu)^3 p(x;\mu,\sigma)dx = 0. \end{aligned}$$

ここで (10) を用いた.

$$\begin{aligned} S(\partial_\mu, \partial_\mu, \partial_\sigma) &= \int_{-\infty}^{\infty} (\partial_\mu \log p(x;\mu,\sigma))^2 (\partial_\sigma \log p(x;\mu,\sigma)) p(x;\mu,\sigma)dx \\ &= -\frac{1}{\sigma^5}\int_{-\infty}^{\infty} (x-\mu)^2 p(x;\mu,\sigma)dx + \frac{1}{\sigma^7}\int_{-\infty}^{\infty} (x-\mu)^4 p(x;\mu,\sigma)dx \\ &= -\frac{1}{\sigma^5}\sigma^2 + \frac{1}{\sigma^7}3\sigma^4 = \frac{2}{\sigma^3} \end{aligned}$$

ここで (9), (11) を用いた.

$$\begin{aligned}
S(\partial_\mu, \partial_\sigma, \partial_\sigma) &= \int_{-\infty}^{\infty} (\partial_\mu \log p(x;\mu,\sigma))(\partial_\sigma \log p(x;\mu,\sigma))^2 p(x;\mu,\sigma)dx \\
&= \frac{1}{\sigma^4}\int_{-\infty}^{\infty}(x-\mu)p(x;\mu,\sigma)dx - \frac{2}{\sigma^6}\int_{-\infty}^{\infty}(x-\mu)^3 p(x;\mu,\sigma)dx \\
&\quad + \frac{1}{\sigma^8}\int_{-\infty}^{\infty}(x-\mu)^5 p(x;\mu,\sigma)dx \\
&= 0.
\end{aligned}$$

ここで (8), (10), (12) を用いた.

$$\begin{aligned}
S(\partial_\sigma, \partial_\sigma, \partial_\sigma) &= \int_{-\infty}^{\infty} (\partial_\sigma \log p(x;\mu,\sigma))^3 p(x;\mu,\sigma)dx \\
&= \frac{1}{\sigma^9}\int_{-\infty}^{\infty}(x-\mu)^6 p(x;\mu,\sigma)dx - \frac{3}{\sigma^7}\int_{-\infty}^{\infty}(x-\mu)^4 p(x;\mu,\sigma)dx \\
&\quad + \frac{3}{\sigma^5}\int_{-\infty}^{\infty}(x-\mu)^2 p(x;\mu,\sigma)dx - \frac{1}{\sigma^3}\int_{-\infty}^{\infty} p(x;\mu,\sigma)dx \\
&= \frac{1}{\sigma^9}15\sigma^6 - \frac{3}{\sigma^7}3\sigma^4 + \frac{3}{\sigma^5}\sigma^2 - \frac{1}{\sigma^3} \\
&= \frac{8}{\sigma^3}
\end{aligned}$$

ここで (9), (11), (13) を用いた.

以上で S の計算は終わった.

3.4 α 接続

以上の結果をもとに, α 接続 $\nabla^{(\alpha)}$ を調べる. ここで α はある実数である.

$$\begin{aligned}
g(\nabla^{(\alpha)}_{\partial_\mu}\partial_\mu, \partial_\mu) &= g(\nabla^{g}_{\partial_\mu}\partial_\mu, \partial_\mu) - \frac{\alpha}{2}S(\partial_\mu, \partial_\mu, \partial_\mu) \\
&= g\left(\frac{1}{2\sigma}\partial_\sigma, \partial_\mu\right) \\
&= 0,
\end{aligned}$$

$$\begin{aligned}
g(\nabla^{(\alpha)}_{\partial_\mu}\partial_\mu, \partial_\sigma) &= g(\nabla^{g}_{\partial_\mu}\partial_\mu, \partial_\sigma) - \frac{\alpha}{2}S(\partial_\mu, \partial_\mu, \partial_\sigma) \\
&= g\left(\frac{1}{2\sigma}\partial_\sigma, \partial_\sigma\right) - \frac{\alpha}{2}\frac{2}{\sigma^3} \\
&= \frac{1-\alpha}{\sigma^3}
\end{aligned}$$

なので,

$$\nabla^{(\alpha)}_{\partial_\mu}\partial_\mu = \frac{1-\alpha}{2\sigma}\partial_\sigma.$$

$$
\begin{aligned}
g(\nabla^{(\alpha)}_{\partial_\mu}\partial_\sigma, \partial_\mu) &= g(\nabla^{g}_{\partial_\mu}\partial_\sigma, \partial_\mu) - \frac{\alpha}{2}S(\partial_\mu, \partial_\sigma, \partial_\mu) \\
&= g\left(-\frac{1}{\sigma}\partial_\mu, \partial_\mu\right) - \frac{\alpha}{2}\frac{2}{\sigma^3} \\
&= -\frac{1+\alpha}{\sigma^3},
\end{aligned}
$$

$$
\begin{aligned}
g(\nabla^{(\alpha)}_{\partial_\mu}\partial_\sigma, \partial_\sigma) &= g(\nabla^{g}_{\partial_\mu}\partial_\sigma, \partial_\sigma) - \frac{\alpha}{2}S(\partial_\mu, \partial_\sigma, \partial_\sigma) \\
&= g\left(-\frac{1}{\sigma}\partial_\mu, \partial_\sigma\right) \\
&= 0
\end{aligned}
$$

なので,

$$
\nabla^{(\alpha)}_{\partial_\mu}\partial_\sigma = -\frac{1+\alpha}{\sigma}\partial_\mu.
$$

$\nabla^{(\alpha)}$ は捩れを持たないので,

$$
\nabla^{(\alpha)}_{\partial_\sigma}\partial_\mu = \nabla^{(\alpha)}_{\partial_\mu}\partial_\sigma = -\frac{1+\alpha}{\sigma}\partial_\mu.
$$

$$
\begin{aligned}
g(\nabla^{(\alpha)}_{\partial_\sigma}\partial_\sigma, \partial_\mu) &= g(\nabla^{g}_{\partial_\sigma}\partial_\sigma, \partial_\mu) - \frac{\alpha}{2}S(\partial_\sigma, \partial_\sigma, \partial_\mu) \\
&= g\left(-\frac{1}{\sigma}\partial_\sigma, \partial_\mu\right) \\
&= 0,
\end{aligned}
$$

$$
\begin{aligned}
g(\nabla^{(\alpha)}_{\partial_\sigma}\partial_\sigma, \partial_\sigma) &= g(\nabla^{g}_{\partial_\sigma}\partial_\sigma, \partial_\sigma) - \frac{\alpha}{2}S(\partial_\sigma, \partial_\sigma, \partial_\sigma) \\
&= g\left(-\frac{1}{\sigma}\partial_\sigma, \partial_\sigma\right) - \frac{\alpha}{2}\frac{8}{\sigma^3} \\
&= -\frac{2+4\alpha}{\sigma^3}
\end{aligned}
$$

なので,

$$
\nabla^{(\alpha)}_{\partial_\sigma}\partial_\sigma = -\frac{1+2\alpha}{\sigma}\partial_\sigma.
$$

以上で $\nabla^{(\alpha)}$ の計算は終わった.

4 曲率テンソルの計算

定曲率統計多様体の定義には様々な流派があるが, この文書では以下の定義で定曲率統計多様体を特徴づけておく.

定義 2: 定曲率統計多様体

統計多様体 (M, g, ∇) が k -定曲率 $(k \in \mathbb{R})$ であるとは,

$$R = kR_0$$

が成り立つことである. ここで R は ∇ の曲率テンソル, R_0 は

$$R_0(X, Y)Z := g(Y, Z)X - g(X, Z)Y$$

で定まるテンソルである.

4.1 R の計算

α 接続 $\nabla^{(\alpha)}$ の曲率テンソル R の計算を行う.

$$\begin{aligned}
R(\partial_\mu, \partial_\sigma)\partial_\mu &= \nabla^{(\alpha)}_{\partial_\mu}\nabla^{(\alpha)}_{\partial_\sigma}\partial_\mu - \nabla^{(\alpha)}_{\partial_\sigma}\nabla^{(\alpha)}_{\partial_\mu}\partial_\mu - \nabla^{(\alpha)}_{[\partial_\mu, \partial_\sigma]}\partial_\mu \\
&= \nabla^{(\alpha)}_{\partial_\mu}\left(-\frac{1+\alpha}{\sigma}\partial_\mu\right) - \nabla^{(\alpha)}_{\partial_\sigma}\left(\frac{1-\alpha}{2\sigma}\partial_\sigma\right) \\
&= -\frac{1+\alpha}{\sigma}\nabla^{(\alpha)}_{\partial_\mu}\partial_\mu + \frac{1-\alpha}{2\sigma^2}\partial_\sigma - \frac{1-\alpha}{2\sigma}\nabla^{(\alpha)}_{\partial_\sigma}\partial_\sigma \\
&= -\frac{1+\alpha}{\sigma}\frac{1-\alpha}{2\sigma}\partial_\sigma + \frac{1-\alpha}{2\sigma^2}\partial_\sigma - \frac{1-\alpha}{2\sigma}\left(-\frac{1+2\alpha}{\sigma}\partial_\sigma\right) \\
&= \frac{1-\alpha^2}{2\sigma^2}\partial_\sigma.
\end{aligned}$$

$$\begin{aligned}
R(\partial_\mu, \partial_\sigma)\partial_\sigma &= \nabla^{(\alpha)}_{\partial_\mu}\nabla^{(\alpha)}_{\partial_\sigma}\partial_\sigma - \nabla^{(\alpha)}_{\partial_\sigma}\nabla^{(\alpha)}_{\partial_\mu}\partial_\sigma - \nabla^{(\alpha)}_{[\partial_\mu, \partial_\sigma]}\partial_\sigma \\
&= \nabla^{(\alpha)}_{\partial_\mu}\left(-\frac{1+2\alpha}{\sigma}\partial_\sigma\right) - \nabla^{(\alpha)}_{\partial_\sigma}\left(-\frac{1+\alpha}{\sigma}\partial_\mu\right) \\
&= -\frac{1+2\alpha}{\sigma}\nabla^{(\alpha)}_{\partial_\mu}\partial_\sigma - \frac{1+\alpha}{\sigma^2}\partial_\mu + \frac{1+\alpha}{\sigma}\nabla^{(\alpha)}_{\partial_\sigma}\partial_\mu \\
&= -\frac{1+2\alpha}{\sigma}\left(-\frac{1+\alpha}{\sigma}\partial_\mu\right) - \frac{1+\alpha}{\sigma^2}\partial_\mu + \frac{1+\alpha}{\sigma}\left(-\frac{1+\alpha}{\sigma}\partial_\mu\right) \\
&= -\frac{1-\alpha^2}{\sigma^2}\partial_\mu.
\end{aligned}$$

以上で R の計算は終わった

4.2 曲率 k の決定

R_0 の計算を行う.

$$R_0(\partial_\mu, \partial_\sigma)\partial_\mu = g(\partial_\sigma, \partial_\mu)\partial_\mu - g(\partial_\mu, \partial_\mu)\partial_\sigma = -\frac{1}{\sigma^2}\partial_\sigma.$$

$$R_0(\partial_\mu, \partial_\sigma)\partial_\sigma = g(\partial_\sigma, \partial_\sigma)\partial_\mu - g(\partial_\mu, \partial_\sigma)\partial_\mu = \frac{2}{\sigma^2}\partial_\mu.$$

以上の計算により,

$$R = -\frac{1-\alpha^2}{2}R_0$$

であることがわかり, $(M, g, \nabla^{(\alpha)})$ が $-\dfrac{1-\alpha^2}{2}$ -定曲率統計多様体であることが示された.

参考文献

[1] 藤原彰夫, 情報幾何学の基礎　情報の内的構造をとらえる新たな地平, 共立出版, 2021.