Pure Mathematics 理论数学, 2025, 15(5), 117-129
Published Online May 2025 in Hans. https://www.hanspub.org/journal/pm
https://doi.org/10.12677/pm.2025.155160

高维正态分布族Fisher度量的曲率

熊明月

重庆理工大学理学院，重庆

收稿日期：2025年3月30日；录用日期：2025年4月30日；发布日期：2025年5月22日

摘　要

本文针对高维情形得到了高维正态分布在Fisher度量下的数量曲率，并且证明了当协方差矩阵Σ为对角矩阵时，正态分布族的参数空间是爱因斯坦空间，其Ricci曲率与度量张量成严格比例关系。

关键词

正态分布，Fisher度量，爱因斯坦空间

Curvature Properties of Fisher Metrics for High-Dimensional Normal Distribution Families

Mingyue Xiong

College of Science, Chongqing University of Technology, Chongqing

Received: Mar. 30th, 2025; accepted: Apr. 30th, 2025; published: May 22nd, 2025

Abstract

In this paper, the scalar curvature of high-dimensional normal distribution under Fisher metric is obtained for the high-dimensional case, and it is proved that when the covariance matrix Σ is a diagonal matrix, the parameter space of the normal distribution family is Einstein space, and its Ricci curvature is strictly proportional to the metric tensor.

Keywords

Normal Distribution, Fisher Metric, Einstein Space

文章引用: 熊明月. 高维正态分布族 Fisher 度量的曲率[J]. 理论数学, 2025, 15(5): 117-129.
DOI: 10.12677/pm.2025.155160

1. 引言

在统计学中，为了对含参数的分布族中的某些参数进行估计，我们需要构造适当的统计量。Fisher 矩阵可以用来刻画统计量的充分性。假设有一个分布族 $p(x,\theta)$，那么 Fisher 矩阵定义为

$$g_{ij}=E\left[\partial_i \log p(x,\theta)\partial_j \log p(x,\theta)\right]$$

由于正态分布作为典型指数族分布，其参数空间的几何特性始终是核心研究议题。正态分布族的 Fisher 矩阵定义了该分布族上的 Riemann 度量。关于这类度量前人已有很多研究工作。早期突破见于 Yoshiharu 等学者(1979) [1]对二维正态分布的探索：他们首次证明当协方差矩阵 Σ 为对角阵时，参数空间具有爱因斯坦流形结构，但其结论限于二维情形，且未揭示高维推广的可能性。Skovgaard (1984) [2]通过将多元正态分布族建模为 Fisher-Rao 度量下的黎曼流形，系统推导了仿射联络与曲率张量的表达式，但其研究止步于黎曼曲率的定性分析，未深入计算数量曲率等全局几何不变量。

本文突破二维情形限制，一方面在 Skovgaard (1984)关于高维正态分布族曲率研究尚不完整的基础上，系统研究了正态分布族的几何结构。得到正态分布族的 Fisher 度量的曲率张量，Ricci 曲率以及数量曲率的完整公式

$$R=-\frac{n(n+1)^2}{4}$$

另一方面得到了当 Σ 为对角矩阵时，正态分布的参数空间是爱因斯坦空间，该结论推广了 Yoshiharu (1979)的二维结果。

2. 预备知识

2.1. 经典信息几何

定义 2.1 [3]：设 $p(x,\theta)$ 是集合 X 上的概率密度函数，其中 θ 是密度函数的参数，称 $S=\{p(x;\theta)\,|\,\theta\in\Theta\}$ 是统计流形，其中 $\Theta\in R^n$。

定义 2.2 [4]：若 n 维统计流形 S 可以表示为：

$$p(x,\theta)=\exp\left\{\theta^i h_i(x)+k(x)-\psi(\theta)\right\}$$

则称 S 是一个指数分布族。其中，x 是 n 维随机变量，θ 是该分布族的 n 维自然坐标，$k(x)$ 是关于 x 的函数，$h_i(x)$ 是 n 个关于 x 线性无关的函数，$\psi(\theta)$ 称为关于参数 θ 的势函数。

若我们规定特定的测量 $\mathrm{d}\mu(x)=\exp\{k(x)\}\mathrm{d}x$，即 $\int p(x,\theta)\mathrm{d}\mu(x)=1$，那么指数族的密度函数就可以写为：

$$p(x,\theta)=\exp\{\theta x-\psi(\theta)\}$$

根据指数分布族的定义，可以将正态分布族写为指数分布族的形式。

首先给出 n 维正态分布的概率密度函数：

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

其中，$\mu^{\mathrm{T}}=(\mu_1,\mu_2,\cdots,\mu_n)\in R^n$ 为正态分布的均值，$\Sigma\in P_n(R)$ 是 n 维协方差矩阵($P_n(R)$ 是 n 维正定对称矩阵)，$x^{\mathrm{T}}=(x_1,x_2,\cdots,x_n)\in R^n$ 为 n 维随机变量。

现在定义出新的随机变量 $x=(x_1,x_2)$：

$$\begin{cases} x_1=h_1(x)=-x \\ x_2=h_2(x)=x^2 \end{cases}$$

同时引入新的参数 (θ,Θ)：

$$\begin{cases} \theta=\Sigma^{-1}\mu \\ \Theta=\frac{1}{2}\Sigma^{-1} \end{cases}$$

那么正态分布族的势函数就可以得到：

$$\psi(\theta)=\frac{1}{4}\theta^{\mathrm{T}}\Theta^{-1}\theta+\frac{n}{2}\log\pi+\frac{1}{2}\log|\Theta|$$

并将正态分布族的 Fisher 信息度量定义为：

$$g_{ij}=E\left[\partial_i\log p(x;\mu,\sigma)\partial_j\log p(x;\mu,\sigma)\right]$$

正态分布族的 Fisher 信息度量是由期望来定义，为了方便计算，在计算过程中一般由 $g_{ij}=-E\left(\partial_i\partial_j\log p(x;\mu,\sigma)\right)$ 进行计算，下面给出它的证明。

定理 2.3 [4]：若 g_{ij} 是光滑的，则

$$E\left(\partial_i\log p(x;\mu,\sigma)\partial_j\log p(x;\mu,\sigma)\right)=-E\left(\partial_i\partial_j\log p(x;\mu,\sigma)\right).$$

2.2. 矩阵的迹

下面介绍关于矩阵以及矩阵迹的相关性质，通过矩阵迹转化为元素的形式，进而得到更为简单的计算。

引理 2.4：若 A,B 是 n 阶矩阵，则

(1)

$$tr\left(AE_{ij}^{*}\right)=\frac{1}{2}\left(a_{ij}+a_{ji}\right) \tag{1}$$

(2) 当 A 或者 B 是 n 阶对称矩阵，有

$$\sum_{i\geq j}tr\left(AE_{ij}^{*}\right)tr\left(BE_{ij}^{*}\right)=\frac{1}{2}tr\left(AB^{\mathrm{T}}\right)+\frac{1}{2}\sum_{i}a_{ii}b_{ii} \tag{2}$$

(3) 当 A 或者 B 是 n 阶对称矩阵，有

$$\sum_{i\geq j}tr\left(AE_{ij}\right)tr\left(BE_{ij}^{*}\right)=tr\left(AB^{\mathrm{T}}\right) \tag{3}$$

(4)

$$\sum_{i\geq j}tr\left(AE_{ij}BE_{ij}^{*}\right)=\frac{1}{2}tr\left(AB^{\mathrm{T}}\right)+\frac{1}{2}tr(A)tr(B) \tag{4}$$

(5) 若 $a=(a_1,a_2,\cdots,a_n)^{\mathrm{T}},b=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$，则

$$\sum_{p}a^{\mathrm{T}}E_{pp}b=a^{\mathrm{T}}b \tag{5}$$

(6) 若 $a=\left(a_1,a_2,\cdots,a_n\right)^{\mathrm{T}}, b=\left(b_1,b_2,\cdots,b_n\right)^{\mathrm{T}}$

$$\sum_p ae^{\mathrm{T}}b=\sum_i b_i a^{\mathrm{T}} \tag{6}$$

(7)

$$\sum_{k\ge l} tr\left(AE_{kl}^*\right)E_{kl}=A \tag{7}$$

(8)

$$\sum_p tr\left(AE_{pp}\right)=tr\left(A\right) \tag{8}$$

(9) 当 A 是 n 阶对称矩阵，有

$$tr\left(AB^{\mathrm{T}}\right)=tr\left(AB\right) \tag{9}$$

其中 $E_{ij}=\begin{cases}1_{i,i} & i=j\\ 1_{i,j}+1_{j,i} & i\ne j\end{cases}$ 且 $E_{ij}^*=\begin{cases}1_{i,i} & i=j\\ \frac{1}{2}\left(1_{i,j}+1_{j,i}\right) & i\ne j\end{cases}$。

证明：

(1)，(8)可以通过简单计算得到，(6)可以将其展开得到，着重证明以下式子。

对于(2)根据公式(2.1)，将迹的形式转换为元素的形式

$$\sum_{i\ge j} tr\left(AE_{ij}^*\right)=\sum_{i\ge j}\frac{1}{2}a_{ij}+\frac{1}{2}a_{ji}$$

$$\sum_{i\ge j} tr\left(AE_{ij}^*\right)tr\left(BE_{ij}^*\right)=\sum_{i\ge j}\left(\frac{1}{4}a_{ij}b_{ij}+\frac{1}{4}a_{ij}b_{ji}+\frac{1}{4}a_{ji}b_{ij}+\frac{1}{4}a_{ji}b_{ji}\right)$$

因为 $A=A^{\mathrm{T}}$，有

$$\sum_{i\ge j} tr\left(AE_{ij}^*\right)tr\left(BE_{ij}^*\right)=\sum_{i\ge j}\left(\frac{1}{2}a_{ij}b_{ji}+\frac{1}{2}a_{ji}b_{ij}\right)=\sum_i a_{ii}b_{ii}+\frac{1}{2}\sum_{i>j}\left(a_{ij}b_{ji}+a_{ji}b_{ij}\right)$$

又因为 $tr\left(AB\right)$ 可以写成

$$tr\left(AB\right)=\sum_{i\ge j}a_{ij}b_{ji}=\sum_i a_{ii}b_{ii}+\sum_{i>j}\left(a_{ij}b_{ji}+a_{ji}b_{ij}\right)$$

所以

$$\sum_{i\ge j} tr\left(AE_{ij}^*\right)tr\left(BE_{ij}^*\right)=\frac{1}{2}tr\left(AB\right)+\frac{1}{2}\sum_i a_{ii}b_{ii}$$

对于(3)， A 或者 B 是 n 阶对称矩阵

$$\sum_{i\ge j} tr\left(AE_{ij}\right)tr\left(BE_{ij}^*\right)=\sum_{i=j}a_{ii}b_{jj}+\sum_{i>j}\left(a_{ij}+a_{ji}\right)\frac{1}{2}\left(b_{ij}+b_{ji}\right)$$

若 A 为对称矩阵就有

$$\sum_{i\ge j} tr\left(AE_{ij}\right)tr\left(BE_{ij}^*\right)=\sum_{i=j}a_{ii}b_{jj}+\sum_{i>j}\left(a_{ij}+a_{ji}\right)\frac{1}{2}\left(b_{ij}+b_{ji}\right)=\sum_{i=j}a_{ii}b_{ii}+\sum_{i>j}a_{ij}\left(b_{ij}+b_{ji}\right)$$

$$tr\left(AB^{\mathrm{T}}\right)=\sum_{i=j}a_{ii}b_{ii}+\sum_{i>j}\left(a_{ij}b_{ij}+a_{ji}b_{ji}\right)=\sum_{i=j}a_{ii}b_{ii}+\sum_{i>j}a_{ij}\left(b_{ij}+b_{ji}\right)$$

将 $\sum_{i\ge j}tr\left(AE_{ij}\right)tr\left(BE_{ij}^*\right)$ 展开，就有

$$tr\left(AB^{\mathrm{T}}\right)=\sum_{i\ge j}tr\left(AE_{ij}\right)tr\left(BE_{ij}^{*}\right)$$

若 B 是对称矩阵就有

$$\sum_{i\ge j}tr\left(AE_{ij}\right)tr\left(BE_{ij}^{*}\right)=\sum_{i=j}a_{ii}b_{jj}+\sum_{i>j}\left(a_{ij}+a_{ji}\right)\frac{1}{2}\left(b_{ij}+b_{ji}\right)=\sum_{i=j}a_{ii}b_{ii}+\sum_{i>j}b_{ij}\left(a_{ij}+a_{ji}\right)$$

综上，若 A 或者 B 是 n 阶对称矩阵，就有

$$tr\left(AB^{T}\right)=\sum_{i\ge j}tr\left(AE_{ij}\right)tr\left(BE_{ij}^{*}\right)$$

对于(4)将 $AE_{ij}BE_{ij}^{*}$ 展开，再计算其迹，有

$$\begin{aligned}\sum_{i\ge j}tr\left(AE_{ij}BE_{ij}^{*}\right)&=\sum_{i}a_{ii}b_{ii}+\frac{1}{2}\sum_{i>j}\left(a_{ij}b_{ij}+a_{ji}b_{ji}+a_{ii}b_{jj}+a_{jj}b_{ii}\right)\\&=\frac{1}{2}\sum_{i>j}\left(a_{ii}b_{ii}+a_{ji}b_{ji}+a_{ij}b_{ij}\right)+\frac{1}{2}\sum_{i>j}\left(a_{ii}b_{ii}+a_{ii}b_{jj}+a_{jj}b_{ii}\right)\\&=\frac{1}{2}tr\left(AB^{\mathrm{T}}\right)+\frac{1}{2}tr\left(A\right)tr\left(B\right)\end{aligned}$$

对于(5)对 $a^{\mathrm{T}}E_{ii}b$ 进行展开，有

$$a^{\mathrm{T}}E_{ii}b=a_ib_i$$

对其求和，所以有

$$\sum_{p}a^{\mathrm{T}}E_{pp}b=\sum_{p}a_pb_p=a^{\mathrm{T}}b$$

对于(7)因为

$$tr\left(AE_{kl}^{*}\right)=\frac{1}{2}\left(a_{kl}+a_{lk}\right)$$

所以有

$$tr\left(AE_{kl}^{*}\right)E_{kl}=\frac{1}{2}\left(a_{kl}+a_{lk}\right)E_{kl}$$

对所有的 k,l 求和即为

$$\sum_{k\ge l}tr\left(AE_{kl}^{*}\right)E_{kl}=\sum_{k\ge l}\frac{1}{2}\left(a_{kl}+a_{lk}\right)E_{kl}=A$$

对于(9)若 A 是 n 阶对称矩阵

$$\begin{aligned}tr\left(AB^{\mathrm{T}}\right)&=tr\left(\left(AB^{\mathrm{T}}\right)^{\mathrm{T}}\right)\\&=tr\left(BA^{\mathrm{T}}\right)\\&=tr\left(BA\right)\\&=tr\left(AB\right)\end{aligned}$$

■

3. 正态分布的 Fisher 信息度量的曲率

在这一部分我们将研究高维正态分布族参数空间的曲率，通常用曲率张量来描述空间的曲率，但要

直观地把握空间的形状，曲率张量尚不充分。因此，我们在这里也要考虑一些二维曲面中由高斯曲率定义的截面曲率。有了前面的计算基础，这章将对高维正态分布族的曲率进行详细的计算。下面介绍后面计算高维正态分布的曲率所要用到的公式

引理 3.1： 设 σ_{ij} 是 Σ 中的元素， μ_i 是 μ 中的元素，那么就有下面等式成立：

(1)

$$\frac{\partial\Sigma}{\partial\sigma_{kl}} = -\Sigma^{-1}E_{kl}\Sigma^{-1} \tag{4.1}$$

(2)

$$\frac{\partial|\Sigma|}{\partial\sigma_{kl}} = |\Sigma| tr\left(\Sigma^{-1}E_{kl}\right) \tag{4.2}$$

(3)

$$\frac{\partial\sigma^{kl}}{\partial\sigma_{ij}} = -tr\left(\Sigma^{-1}E_{ij}\Sigma^{-1}E_{kl}^{*}\right) \tag{4.3}$$

(4)

$$E\left[\left(x-\mu\right)^{\mathrm{T}} A\left(x-\mu\right)\right] = tr\left(A\Sigma\right) \tag{4.4}$$

(5)

$$\frac{\partial tr\left(A^{-1}E_{kl}\right)}{\partial a_{ij}} = -tr\left(A^{-1}E_{ij}A^{-1}E_{kl}\right) \tag{4.5}$$

证明： 对于(1)，因为

$$\Sigma\Sigma^{-1} = E$$

对两边求 σ_{kl} 偏导有

$$\begin{aligned}\frac{\partial\left(\Sigma\Sigma^{-1}\right)}{\partial\sigma_{kl}} &= \mathbf{0}\\ &= \frac{\partial\Sigma}{\partial\sigma_{kl}}\Sigma^{-1} + \frac{\partial\Sigma^{-1}}{\partial\sigma_{kl}}\Sigma\\ &= E_{kl}\Sigma^{-1} + \frac{\partial\Sigma^{-1}}{\partial\sigma_{kl}}\Sigma\end{aligned}$$

化简过后得

$$\frac{\partial\Sigma}{\partial\sigma_{kl}} = -\Sigma^{-1}E_{kl}\Sigma^{-1}$$

对于(2)，设 $A^{*} = B$

有 $|A|I = A^{*}A = BA$，对其矩阵展开并取出 $|A|I$ 的 k 行 k 列的元素，就有

$$|A| = \sum_{p} b_{kp}a_{pk}$$

对其两边求 a_{kl} 的偏导

$$\frac{\partial|A|}{a_{kl}} = \begin{cases} b_{kl} + b_{lk} & k \neq l \\ b_{kk} & k = l \end{cases}$$

即

$$\frac{\partial|A|}{\partial a_{kl}}=tr\left(BE_{kl}\right)=|A|tr\left(A^{-1}E_{kl}\right)$$

对于(3)，因为Σ^{-1}是对称矩阵，与$\sigma^{kl}=tr\left(\Sigma^{-1}E_{kl}\right)$，对$\sigma^{kl}$求偏导有

$$\frac{\partial\sigma^{kl}}{\partial\sigma_{ij}}=\frac{\partial}{\partial\sigma_{ij}}tr\left(\Sigma^{-1}E_{kl}\right)=\frac{\partial}{\partial\sigma_{ij}}\left(\Sigma^{-1}\right)_{kl}=\left(\frac{\partial\Sigma^{-1}}{\partial\sigma_{ij}}\right)_{kl}=-\left(\Sigma^{-1}E_{ij}\Sigma^{-1}\right)_{kl}$$

因为$\Sigma^{-1}E_{ij}\Sigma^{-1}$是对称矩阵，所以有

$$\frac{\partial\sigma^{kl}}{\partial\sigma_{ij}}=-tr\left(\Sigma^{-1}E_{ij}\Sigma^{-1}E_{kl}^{*}\right)$$

对于(4),将$(x-\mu)^{\mathrm{T}}A(x-\mu)$展开，得到

$$(x-\mu)^{\mathrm{T}}A(x-\mu)=\sum_{i,j}a_{ij}\left(x_i-\mu_i\right)\left(x_j-\mu_j\right)$$

对其求期望，所以有

$$\begin{aligned}E\left[(x-\mu)^{\mathrm{T}}A(x-\mu)\right]&=\sum_{i,j}a_{ij}\left(E\left(x_ix_j\right)-E\left(x_i\right)\mu_j-E\left(x_j\right)\mu_i-\mu_i\mu_j\right)\\&=\sum_{i,j}a_{ij}\sigma_{ij}\\&=tr\left(A\Sigma\right)\end{aligned}$$

对于(5)，

$$\frac{\partial tr\left(A^{-1}E_{kl}\right)}{\partial a_{ij}}=2\left(\frac{\partial A^{-1}}{\partial a_{ij}}\right)_{kl}=-2\left(A^{-1}E_{ij}A^{-1}\right)_{kl}=-2tr\left(A^{-1}E_{ij}A^{-1}E_{kl}^{*}\right)=-tr\left(A^{-1}E_{ij}A^{-1}E_{kl}\right)$$

■

定理 3.2 [2]：设g是高维正态分布的信息度量，则其度量的张量为

$$\begin{cases}g_{i,kl}=0\\g_{ij,kl}=\dfrac{1}{2}tr\left(\Sigma^{-1}E_{ij}\Sigma^{-1}E_{kl}\right)\\g_{i,j}=\sigma^{ij}\end{cases}$$

度量g的逆矩阵就可以表示为

$$\begin{cases}g^{i,kl}=0\\g^{ij,kl}=2tr\left(\Sigma E_{ij}^{*}\Sigma E_{kl}^{*}\right)\\g^{i,j}=\sigma_{ij}\end{cases}$$

定理 3.3 [2]：正态分布的联络系数可表示为

$$\begin{cases}\Gamma_{i,j}^{k}=\Gamma_{ij,k}^{rs}=\Gamma_{ij,kl}^{r}=0\\\Gamma_{i,j}^{kl}=tr\left(E_{kl}^{*}E_{ij}^{*}\right)\\\Gamma_{ij,k}^{l}=-\dfrac{1}{2}e_k^{\mathrm{T}}\Sigma^{-1}E_{ij}e_l\\\Gamma_{ij,kl}^{rs}=-\dfrac{1}{2}tr\left(E_{rs}^{*}E_{kl}\Sigma^{-1}E_{ij}\right)-\dfrac{1}{2}tr\left(E_{rs}^{*}E_{ij}\Sigma^{-1}E_{kl}\right)\end{cases}$$

定理 3.4 [2]：若 $e_i \in R^n$ 表示 μ -方向的基向量场， $E_{ij} \in R^{n*n}$ 表示 σ -方向的基向量场，那么正态分布族的仿射联络 ∇ 可以表示为

$$\begin{cases} \nabla_{e_i} e_j = \nabla_{e_j} e_i = \frac{1}{2}\left(e_i e_j^{\mathrm{T}} + e_j e_i^{\mathrm{T}}\right) \\ \nabla_{e_i} E_{kl} = \nabla_{E_{kl}} e_i = \frac{1}{2} E_{kl} \Sigma^{-1} e_i \\ \nabla_{E_{ij}} E_{kl} = \nabla_{E_{kl}} E_{ij} = \frac{1}{2}\left(E_{kl}\Sigma^{-1}E_{ij} + E_{ij}\Sigma^{-1}E_{kl}\right) \end{cases}$$

定理 3.5 [2]：黎曼曲率张量可以表示为

$$\begin{cases} R_{i,j,k,l} = -\frac{1}{4}\sigma^{ki}\sigma^{jl} + \frac{1}{4}\sigma^{kj}\sigma^{il} \\ R_{ij,kl,rs,pq} = \frac{1}{4} tr\left(\Sigma^{-1}E_{kl}\Sigma^{-1}E_{ij}\Sigma^{-1}E_{pq}\Sigma^{-1}E_{rs}\right) - \frac{1}{4} tr\left(\Sigma^{-1}E_{ij}\Sigma^{-1}E_{kl}\Sigma^{-1}E_{pq}\Sigma^{-1}E_{rs}\right) \\ R_{i,j,kl,rs} = \frac{1}{4}\left(e_i^{\mathrm{T}}\Sigma^{-1}E_{kl}\Sigma^{-1}E_{rs}\Sigma^{-1}e_j - e_i^{\mathrm{T}}\Sigma^{-1}E_{rs}\Sigma^{-1}E_{kl}\Sigma^{-1}e_j\right) \\ R_{i,kl,j,rs} = \frac{1}{4} e_i^{\mathrm{T}}\Sigma^{-1}E_{kl}\Sigma^{-1}E_{rs}\Sigma^{-1}e_j \end{cases}$$

下面我们介绍如何利用黎曼曲率张量计算 n 维正态分布的数量曲率。

定理 3.6：n 维正态分布的数量曲率为

$$R = -\frac{n(n+1)^2}{4}$$

证明：

首先计算 Ricci 张量的分量：

$$\sum_{k,l} R_{k,i,j,l} g^{k,l} = \frac{1}{4}\sum_{k,l}\left(\sigma^{kl}\sigma^{ij} - \sigma^{kj}\sigma^{il}\right)\sigma_{kl} = \frac{n-1}{4}\sigma^{ij}$$

$$\begin{aligned} \sum_{\substack{k\ge l\\ p\ge q}} R_{kl,i,j,pq} g^{kl,pq} &= -\frac{1}{2}\sum_{\substack{k\ge l\\ p\ge q}} e_i^{\mathrm{T}}\Sigma^{-1}E_{kl}\Sigma^{-1}E_{pq}\Sigma^{-1}e_j tr\left(\Sigma E_{kl}^{*}\Sigma E_{pq}^{*}\right) \\ &= \frac{1}{2}\sum_{\substack{k\ge l\\ p\ge q}} tr\left(e_i^{\mathrm{T}}\Sigma^{-1}E_{kl}\Sigma^{-1}E_{pq}\Sigma^{-1}e_j\right) tr\left(\Sigma E_{kl}^{*}\Sigma E_{pq}^{*}\right) \\ &= -\frac{1}{2}\sum_{\substack{k\ge l\\ p\ge q}} tr\left(\Sigma^{-1}e_j e_i^{\mathrm{T}}\Sigma^{-1}E_{kl}\Sigma^{-1}E_{pq}\right) tr\left(\Sigma E_{kl}^{*}\Sigma E_{pq}^{*}\right) \end{aligned}$$

由公式(3)

$$\begin{aligned} \sum_{\substack{k\ge l\\ p\ge q}} R_{kl,i,j,pq} g^{kl,pq} &= -\frac{1}{2}\sum_{\substack{k\ge l\\ p\ge q}} tr\left(\Sigma^{-1}e_j e_i^{\mathrm{T}}\Sigma^{-1}E_{kl}\Sigma^{-1}\Sigma E_{kl}^{*}\Sigma\right) \\ &= -\frac{1}{2}\sum_{\substack{k\ge l\\ p\ge q}} tr\left(e_j e_i^{\mathrm{T}}\Sigma^{-1}E_{kl}E_{kl}^{*}\right) \end{aligned}$$

由公式(4)

$$\sum_{\substack{k\ge l\\ p\ge q}} R_{kl,i,j,pq} g^{kl,pq} = -\frac{1}{4}\left(tr\left(e_j e_i^{\mathrm{T}}\Sigma^{-1}I\right) + tr\left(e_j e_i^{\mathrm{T}}\Sigma^{-1}\right) tr(I)\right) = -\frac{n+1}{4}\sigma^{ij}$$

计算 $\sum_{p,q} R_{p,ij,kl,q} g^{p,q}$

$$\sum_{p,q} R_{p,ij,kl,q} g^{p,q} = -\frac{1}{4} e_p^{\mathrm{T}} \Sigma^{-1} E_{ij} \Sigma^{-1} E_{kl} \Sigma^{-1} e_q \sigma_{pq}$$
$$= -\frac{1}{4} e_p^{\mathrm{T}} \Sigma^{-1} E_{ij} \Sigma^{-1} E_{kl} \Sigma^{-1} e_q e_q^{T} \Sigma e_p$$
$$= -\frac{1}{4} e_p^{\mathrm{T}} \Sigma^{-1} E_{ij} \Sigma^{-1} E_{kl} \Sigma^{-1} E_{qq} \Sigma e_p$$

由公式(8)，可得

$$\sum_{p,q} R_{p,ij,kl,q} g^{p,q} = -\frac{1}{4} e_p^{\mathrm{T}} \Sigma^{-1} E_{ij} \Sigma^{-1} E_{kl} e_p$$
$$= -\frac{1}{4} tr\left(e_p^{\mathrm{T}} \Sigma^{-1} E_{ij} \Sigma^{-1} E_{kl} e_p \right)$$
$$= -\frac{1}{4} tr\left(\Sigma^{-1} E_{ij} \Sigma^{-1} E_{kl} E_{pp} \right)$$

由公式(3)，可得

$$\sum_{p,q} R_{p,ij,kl,q} g^{p,q} = -\frac{1}{4} tr\left(\Sigma^{-1} E_{ij} \Sigma^{-1} E_{kl} \right)$$

下面计算

$$\sum_{\substack{p\ge q\\ k\ge l}} R_{pq,ij,kl,rs} g^{pq,rs} = \frac{1}{4} tr\left(\Sigma^{-1} E_{rs} \Sigma^{-1} E_{kl} \Sigma^{-1} E_{pq} \Sigma^{-1} E_{ij} \right) 2tr\left(\Sigma E_{rs}^{*} \Sigma E_{pq}^{*} \right)$$
$$-\frac{1}{4} tr\left(\Sigma^{-1} E_{rs} \Sigma^{-1} E_{kl} \Sigma^{-1} E_{ij} \Sigma^{-1} E_{pq} \right) 2tr\left(\Sigma E_{rs}^{*} \Sigma E_{pq}^{*} \right)$$

由公式(3)，可得

$$\sum_{\substack{p\ge q\\ k\ge l}} R_{pq,ij,kl,rs} g^{pq,rs} = \frac{1}{2} tr\left(E_{ij} \Sigma^{-1} E_{rs} \Sigma^{-1} E_{kl}\, E_{rs}^{*} \right) - \frac{1}{2} tr\left(E_{rs} \Sigma^{-1} E_{kl} \Sigma^{-1} E_{ij} E_{rs}^{*} \right)$$

由公式(4)，可得

$$\sum_{\substack{p\ge q\\ k\ge l}} R_{pq,ij,kl,rs} g^{pq,rs} = \frac{1}{4} tr\left(E_{ij} \Sigma^{-1} E_{kl} \Sigma^{-1} \right) + \frac{1}{4} tr\left(E_{ij} \Sigma^{-1} \right) tr\left(E_{kl} \Sigma^{-1} \right)$$
$$-\frac{1}{4} tr\left(E_{ij} \Sigma^{-1} E_{kl} \Sigma^{-1} \right) - \frac{1}{4} tr\left(E_{ij} \Sigma^{-1} E_{kl} \Sigma^{-1} \right) tr\left(I \right)$$
$$= \frac{1}{4} tr\left(E_{ij} \Sigma^{-1} \right) tr\left(E_{kl} \Sigma^{-1} \right) - \frac{n}{4} tr\left(E_{ij} \Sigma^{-1} E_{kl} \Sigma^{-1} \right)$$

对于二维正态分布的参数空间，Ricci 张量的分量由下面的公式给出：

$$R_{ij} = \sum_{k,l} R_{i,j,k,l} g^{k,l} + \sum_{\substack{k\ge l\\ p\ge q}} R_{kl,i,j,pq} g^{kl,pq}$$
$$= \frac{n-1}{4} \sigma^{ij} - \frac{n+1}{4} \sigma^{ij}$$
$$= -\frac{1}{2} \sigma^{ij}$$

$$
\begin{aligned}
R_{ij,kl} &= \sum_{p,q} R_{p,ij,kl,q} g^{p,q} + \sum_{\substack{k\ge l\\ p\ge q}} R_{pq,ij,kl,rs} g^{pq,rs} \\
&= -\frac{1}{4} tr\left(\Sigma^{-1}E_{ij}\Sigma^{-1}E_{kl}\right) + \frac{1}{4} tr\left(E_{ij}\Sigma^{-1}\right) tr\left(E_{kl}\Sigma^{-1}\right) - \frac{n}{4} tr\left(E_{ij}\Sigma^{-1}E_{kl}\Sigma^{-1}\right) \\
&= -\frac{n+1}{4} tr\left(\Sigma^{-1}E_{ij}\Sigma^{-1}E_{kl}\right) + \frac{1}{4} tr\left(E_{ij}\Sigma^{-1}\right) tr\left(E_{kl}\Sigma^{-1}\right)
\end{aligned}
$$

那么数量曲率可以表示为

$$
\begin{aligned}
R &= \sum_{i,j} R_{i,j} g^{i,j} + \sum_{\substack{i\ge j\\ k\ge l}} R_{ij,kl} g^{ij,kl} \\
&= -\frac{1}{2}\sigma^{ij}\sigma_{ij} - \frac{n+1}{4} tr\left(\Sigma^{-1}E_{ij}\Sigma^{-1}E_{kl}\right) * 2tr\left(\Sigma E_{ij}^{*}\Sigma E_{kl}^{*}\right) \\
&\quad + \frac{1}{4} tr\left(E_{ij}\Sigma^{-1}\right) tr\left(E_{kl}\Sigma^{-1}\right) * 2tr\left(\Sigma E_{ij}^{*}\Sigma E_{kl}^{*}\right)
\end{aligned}
$$

由公式(3)，可得

$$
\begin{aligned}
R &= -\frac{n}{2} - \frac{n+1}{2}\sum_{i\ge j} tr\left(E_{ij}E_{ij}^{*}\right) + \frac{1}{2} tr\left(E_{ij}\Sigma^{-1}\right) tr\left(\Sigma E_{ij}^{*}\right) \\
&= -\frac{n}{2} - \frac{n+1}{2} * \frac{n(n+1)}{2} + \frac{n}{2} \\
&= -\frac{n(n+1)^2}{4}
\end{aligned}
$$

■

上述定理可知 n 维正态分布的数量曲率是跟维数有关的常数，但是通过 Ricci 张量和黎曼张量的比较，我们发现一般情况下我们的参数空间不是爱因斯坦空间，但是我们发现在某个特定的子流形下，高维正态分布的参数空间是爱因斯坦空间。下面给出高维正态分布的参数空间是爱因斯坦空间的充分条件。

推论 3.7： 若 Σ 是对角矩阵，则 n-维正态分布的参数空间是爱因斯坦空间。

证明： 当 Σ 是对角矩阵时，根据度量公式

$$
g_{ij} = -E\left[\partial_i\partial_j p(x;\mu,\Sigma)\right]
$$

可得到正态分布的黎曼度量为

$$
g = \begin{cases}
g_{i,jj} = 0; \\
g_{ii,jj} = \begin{cases} 0, & ii \ne jj \\ \frac{1}{2}\left(\sigma^{ii}\right)^2, & ii = jj \end{cases}; \\
g_{i,j} = \sigma^{ij}
\end{cases}
$$

因为 $g_{ii,jj}$ 和 $g_{i,j}$ 是对称矩阵，所以有正态分布的 Fisher 度量的逆矩阵

$$
g^{-1} = \begin{cases}
g^{i,jj} = 0; \\
g^{ii,jj} = \begin{cases} 0, & ii \ne jj \\ 2\sigma_{ii}^2, & ii = jj \end{cases}; \\
g^{i,j} = \sigma_{ij}
\end{cases}
$$

现在计算联络系数

$$\Gamma_{i,j}^{kk} = -\frac{1}{2} g^{kk,kk} \frac{\partial g_{i,i}}{\partial \sigma_{kk}} = tr\left(E_{kk} E_{ii}\right)$$

当且仅当 $k = i$ 时， $tr\left(E_{kk} E_{ii}\right) = 1$，其余为零，即

$$\Gamma_{k,k}^{kk} = 1$$

$$\Gamma_{ii,k}^{l} = -\frac{1}{2} e_k^{\mathrm{T}} \Sigma^{-1} E_{ii} e_l$$

由于矩阵 $\Sigma^{-1} E_{ii}$ 是对称矩阵所以有

$$\Gamma_{ii,k}^{l} = -\frac{1}{2} tr\left(\Sigma^{-1} E_{ii} E_{kl}\right)$$

这个式子只有当 $k = i = l$ 时不为零，即

$$\Gamma_{ii,i}^{i} = -\frac{1}{2} tr\left(\Sigma^{-1} E_{ii} E_{ii}\right) = -\frac{1}{2} \sigma^{ii}$$

$$\Gamma_{ii,kk}^{rr} = -\frac{1}{2} tr\left(E_{rr} E_{kk} \Sigma^{-1} E_{ii}\right) - \frac{1}{2} tr\left(E_{rr} E_{ii} \Sigma^{-1} E_{kk}\right)$$

这个式子只有当 $k = i = r$ 时不为零，即

$$\Gamma_{ii,ii}^{ii} = -\frac{1}{2} \sigma^{ii} - \frac{1}{2} \sigma^{ii} = -\sigma^{ii}$$

$$\begin{cases} \Gamma_{k,k}^{kk} = 1 \\ \Gamma_{ii,i}^{i} = -\frac{1}{2} \sigma^{ii} \\ \Gamma_{ii,ii}^{ii} = -\sigma^{ii} \end{cases}$$

其余的联络系数为 0。

正态分布族的仿射联络 ∇ 为

$$\nabla_{e_i} e_i = \Gamma_{i,i}^{ii} E_{ii} = E_{ii}$$

$$\nabla_{ei} E_{ii} = \Gamma_{i,ii}^{i} e_i = -\frac{1}{2} \sigma^{ii} e_i$$

$$\nabla_{E_{ii}} E_{ii} = \Gamma_{ii,ii}^{ii} E_{ii} = -\sigma^{ii} E_{ii}$$

$$\begin{cases} \nabla_{e_i} e_i = E_{ii} \\ \nabla_{e_i} E_{ii} = -\frac{1}{2} \sigma^{ii} e_i \\ \nabla_{E_{ii}} E_{ii} = -\sigma^{ii} E_{ii} \end{cases}$$

其余为 0。

接下来求二阶偏导：

$$\begin{aligned} \nabla_{e_i} \nabla_{e_i} e_i &= \nabla_{e_i} \left(\nabla_{e_i} e_i\right) \\ &= \nabla_{e_i} E_{ii} \\ &= -\frac{1}{2} \sigma^{ii} e_i \end{aligned}$$

$$\begin{aligned}\nabla_{E_{ii}}\nabla_{E_{ii}}E_{ii} &= \nabla_{E_{ii}}\left(\nabla_{E_{ii}}E_{ii}\right)\\ &= -\sigma^{ii}\nabla_{e_i}E_{ii}\\ &= -\left(\sigma^{ii}\right)^2 E_{ii}\end{aligned}$$

$$\begin{aligned}\nabla_{e_i}\nabla_{e_i}E_{ii} &= \nabla_{e_i}\left(\nabla_{e_i}E_{ii}\right)\\ &= -\frac{1}{2}\sigma^{ii}\nabla_{e_i}e_i\\ &= -\frac{1}{2}\sigma^{ii}E_{ii}\end{aligned}$$

$$\nabla_{E_{ii}}\nabla_{e_i}e_i = \nabla_{E_{ii}}E_{ii} = -\sigma^{ii}E_{ii}$$

$$\begin{cases}\nabla_{e_i}\nabla_{e_i}e_i = -\frac{1}{2}\sigma^{ii}e_i\\ \nabla_{E_{ii}}\nabla_{E_{ii}}E_{ii} = -\left(\sigma^{ii}\right)^2 E_{ii}\\ \nabla_{e_i}\nabla_{e_i}E_{ii} = -\frac{1}{2}\sigma^{ii}E_{ii}\\ \nabla_{E_{ii}}\nabla_{e_i}e_i = -\sigma^{ii}E_{ii}\end{cases}$$

其余为 0。

黎曼曲率张量就可以表示为

$$\begin{aligned}R_{i,ii,i,ii} &= g\left(\nabla_{e_i}\nabla_{E_{ii}}e_i - \nabla_{E_{ii}}\nabla_{e_i}e_i, E_{ii}\right)\\ &= g\left(-\frac{1}{2}\sigma^{ii}E_{ii} + \sigma^{ii}E_{ii}, E_{ii}\right)\\ &= \frac{1}{2}\sigma^{ii}g\left(E_{ii}, E_{ii}\right)\\ &= \frac{1}{4}\left(\sigma^{ii}\right)^3\end{aligned}$$

其余为 0。

Ricci 曲率表示为

$$\begin{aligned}R_{i,i} &= R_{ii,i,i,ii}g^{ii,ii}\\ &= -R_{i,ii,i,ii}g^{ii,ii}\\ &= -\frac{1}{4}\left(\sigma^{ii}\right)^3 2\sigma_{ii}^2\\ &= -\frac{1}{2}\sigma^{ii}\end{aligned}$$

$$\begin{aligned}R_{ii,ii} &= R_{i,ii,ii,i}g^{i,i}\\ &= -R_{i,ii,i,ii}g^{i,i}\\ &= -\frac{1}{4}\left(\sigma^{ii}\right)^3\sigma^{ii}\\ &= -\frac{1}{4}\left(\sigma^{ii}\right)^2\end{aligned}$$

$$\begin{cases} R_{i,i} = -\frac{1}{2}\sigma^{ii} \\ R_{ii,ii} = -\frac{1}{4}\left(\sigma^{ii}\right)^2 \end{cases}$$

这个空间的数量曲率为

$$\begin{aligned} R &= \sum_i R_{i,i} g^{i,i} + R_{ii,ii} g^{ii,ii} \\ &= \sum_i \left(-\frac{1}{2}\sigma^{ii}\sigma_{ii} - \frac{1}{4}\left(\sigma^{ii}\right)^2 2\sigma_{ii}^2 \right) \\ &= -n \end{aligned}$$

比较度量张量与 Ricci 张量，有如下的关系：

$$\begin{cases} R_{i,i} = -\frac{R}{2n} g_{ii} = -\frac{1}{2} g_{ii} \\ R_{ii,ii} = -\frac{R}{2n} g_{ii,ii} = -\frac{1}{2} g_{ii,ii} \end{cases}$$

综上可以看到，n-维正态分布的参数空间是爱因斯坦空间。

■

参考文献

[1] Sato, Y., Sugawa, K. and Kawaguchi, M. (1979) The Geometrical Structure of the Parameter Space of the Two-Dimensional Normal Distribution. *Reports on Mathematical Physics*, **16**, 111-119. https://doi.org/10.1016/0034-4877(79)90043-0

[2] Skovgaard, L.T. (1984) A Riemannian Geometry of the Multivariate Normal Model. *Scandinavian Journal of Statistics*, **11**, 211-223.

[3] 孙华飞, 张真宁, 彭林玉, 段晓敏. 信息几何导引[M]. 北京: 科学出版社, 2016.

[4] Amari, S. (2016) Information Geometry and Its Applications. Springer.