α-SCALAR CURVATURE OF THE t-MANIFOLD

Bong Sik Cho* and Sun Young Jung

Abstract. The Fisher information matrix plays a significant role in statistical inference in connection with estimation and properties of variance of estimators. In this paper, we define the parameter space of the t-manifold using its Fisher's matrix and characterize the t-manifold from the viewpoint of information geometry. The α-scalar curvatures to the t-manifold are calculated.

1. Introduction

Information geometry is the differential geometric study of the manifold of probability measures or probability density functions. Recently, information geometric methods have been applied to many areas of the study of estimating functions and nuisance parameter, the dependency of Bayesian predictive distribution, the class of invariant priors for Bayesian inference, principal component analysis, independent component analysis and blind source separation.

Rao (1945) first noticed the importance of the differential-geometrical approach and introduced the Riemannian metric in a statistical manifold by using the Fisher information matrix and calculated the geodesic distance between two distributions for various statistical models. Since then many researchers have tried to obtain the properties of the Riemannian manifold of a statistical model. Efron (1975) elucidated the meaning of curvature for asymptotic statistical inference and pointed that the statistical curvature plays a fundamental role in the higher order asymptotic theory of statistical inference. Amari (1982) gave a natural definition of a family of Affine connections on the statistical manifolds,

so called the α-connection and α-curvature. Then he pointed out important roles of the exponential and mixture curvatures and their duality in statistical inference. Amari (1985) remarked that the two dimensional parameter space of the family of one dimensional normal distribution is a space of negative constant curvature and studied the α-geometry of the families of the gamma, Gaussian, Mckey bivariate gamma and the Freund bivariate exponential. Recently, Adbel-All et al. (2003), Kass (1989), Kass and Vos (1997), Murray and Rice (1993) studied the probability density function from the viewpoint of information geometry and use the geometric metrics to give a new description to the statistical distribution. Arwini and Dodson (2007) studied the α-geometry of the Weibull manifold. In this paper, we find the α-connection and α-scalar curvature of the t-manifold.

2. The t-manifold

The set
$$S = \{p(x) = \frac{1}{v}\frac{\Gamma(\frac{r+1}{2})}{\sqrt{\pi r}\Gamma(\frac{r}{2})}(1 + \frac{1}{r}(\frac{x-u}{v})^2)^{-\frac{r+1}{2}} | x \in R, (u,v) \in R \times R^+\}$$
is called the t-manifold, where

$$(2.1) \qquad p(x) = \frac{1}{v}\frac{\Gamma(\frac{r+1}{2})}{\sqrt{\pi r}\Gamma(\frac{r}{2})}(1 + \frac{1}{r}(\frac{x-u}{v})^2)^{-\frac{r+1}{2}}$$

is the probability density function of the t-distribution location parameter.
Define

$$a = \frac{1}{r}, \; b = \frac{r+1}{2}, \; c_r = \frac{\Gamma(\frac{r+1}{2})}{\sqrt{\pi r}\Gamma(\frac{r}{2})}$$

From (2.1), the log likelihood function is

$$(2.2) \qquad \ln p(x) = \ln c_r - b\ln(1 + a(\frac{x-u}{v})^2) - \ln v$$

Set $l(x) = \ln p(x)$. Taking the coordinate $(\theta_1, \theta_2) = (u, v)$ and setting $\partial_i l = \frac{\partial}{\partial \theta_i}l(x)$, from (2.2) we get

$$\partial_1\partial_1 l = \frac{2ab(a(x-u)^2 - v^2)}{(v^2 + a(x-u)^2)^2}, \; \partial_1\partial_2 l = \frac{4abv(x-u)}{(v^2 + a(x-u)^2)^2}$$

$$\partial_2\partial_2 l = \frac{1}{v^2}[1 + \frac{-6ab(x-u)^2v^{-2}(1+a(x-u)^2v^{-2}) + 4a^2b(x-u)^4v^{-4}}{(1+a(x-u)^2v^{-2})^2}].$$

From Cho and Baek(2006), Fisher information matrix (g_{ij}) is given by

$$(2.3) \quad (g_{ij}) = \begin{pmatrix} \frac{r+1}{v^2(r+3)} & 0 \\ 0 & \frac{2r}{v^2(r+3)} \end{pmatrix}, \quad (g^{ij}) = \begin{pmatrix} \frac{v^2(r+3)}{r+1} & 0 \\ 0 & \frac{v^2(r+3)}{2r} \end{pmatrix}$$

We get

$$ds^2 = \frac{1}{v^2(r+3)}((r+1)du^2+2rdv^2), \ dA = \sqrt{\det(g_{ij})} = \frac{\sqrt{2r(r+1)}}{v^2(r+3)}dudv.$$

Moreover

$$(2.4) \quad E(\frac{1}{Z^2+1}) = \frac{r}{r+1}, \ E(\frac{1}{(Z^2+1)^2}) = \frac{r(r+2)}{(r+1)(r+3)},$$

$$(2.5) \quad E(\frac{1}{(Z^2+1)^3}) = \frac{r(r+2)(r+4)}{(r+1)(r+3)(r+5)}$$

Proposition 2.1. *The α-curvature tensor of the t-manifold is given by*

$$(2.6)$$
$$R^{(\alpha)}_{1212} = -\frac{(r+1)\{12(r+2) + 2(1-\alpha)(r-1)^2 - (1-\alpha)^2(r-1)^2\}}{v^4(r+3)(r+5)^2}.$$

Proof. Since the α-connection is defined by

$$(2.7) \quad \begin{aligned} \Gamma^{(\alpha)}_{ijk} &= E[(\partial_i\partial_j l + \frac{1-\alpha}{2}\partial_i l\partial_j l)\partial_k l] \\ &= \Gamma^{(1)}_{ijk} + \frac{1-\alpha}{2}T_{ijk}, \end{aligned}$$

where $T_{ijk} = E[\partial_i l\partial_j l\partial_k l]$, from (2.4) and (2.5)

(2.8)

$$\Gamma_{111}^{(1)} = E[\partial_1\partial_1 l \partial_1 l] = \frac{4a\sqrt{a}b^2}{v^3}E[\frac{Z(Z^2-1)}{(1+Z^2)^3}] = 0$$

$$\Gamma_{122}^{(1)} = E[\partial_1\partial_2 l \partial_2 l] = \frac{4\sqrt{a}b}{v^3}E[\frac{Z^3(2b-1)-Z}{(1+Z^2)^3}] = 0$$

$$\Gamma_{112}^{(1)} = E[\partial_1\partial_1 l \partial_2 l] = \frac{2ab}{v^3}E[\frac{(2b-1)Z^4-2bZ^2+1}{(1+Z^2)^3}] = \frac{6(r+1)}{v^3(r+3)(r+5)}$$

$$\Gamma_{121}^{(1)} = E[\partial_1\partial_2 l \partial_1 l] = -\frac{8ab^2}{v^3}E[\frac{Z^2}{(1+Z^2)^3}] = -\frac{2(r+1)(r+2)}{v^3(r+3)(r+5)}$$

$$\Gamma_{222}^{(1)} = E[\partial_2\partial_2 l \partial_2 l] = -\frac{1}{v^3}E[\frac{r^2(Z^6+3Z^4)-(4r+1)Z^2+1}{(1+Z^2)^3}]$$

$$= -\frac{6r(r+1)}{v^3(r+3)(r+5)}$$

where $a = 1/r$ and $b = \dfrac{r+1}{2}$.

By (2.7)

(2.9)

$$T_{112} = E[(\partial_1 l)^2 \partial_2 l] = \frac{4ab^2}{v^3}\{(2b-1)E[\frac{Z^4}{(1+Z^2)^3}] - E[\frac{Z^2}{(1+Z^2)^3}]\}$$

$$= \frac{2(r+1)(r-1)}{v^3(r+3)(r+5)}$$

$$T_{222} = E[(\partial_2 l)^3] = \frac{1}{v^3}\{(2b-1)^3 - 6b(2b-1)^2 E[\frac{1}{1+Z^2}]$$

$$+ 12b^2(2b-1)E[\frac{1}{(1+Z^2)^2}] - 8b^3 E[\frac{1}{(1+Z^2)^3}]\} = \frac{8r(r-1)}{v^3(r+3)(r+5)}$$

$$T_{111} = T_{122} = 0,$$

From (2.7), (2.8) and (2.9)

(2.10)

$$\Gamma_{112}^{(\alpha)} = \Gamma_{112}^{(1)} + \frac{1-\alpha}{2} T_{112} = \frac{6(r+1) + (1-\alpha)(r+1)(r-1)}{v^3(r+3)(r+5)}$$

$$\Gamma_{222}^{(\alpha)} = \Gamma_{222}^{(1)} + \frac{1-\alpha}{2} T_{222} = \frac{-6r(r+1) + (1-\alpha)4r(r-a)}{v^3(r+3)(r+5)}$$

$$\Gamma_{121}^{(\alpha)} = \Gamma_{121}^{(1)} + \frac{1-\alpha}{2} T_{121} = \frac{-2(r+1)(r+2) + (1-\alpha)(r+1)(r-1)}{v^3(r+3)(r+5)}$$

$$\Gamma_{111}^{(\alpha)} = \Gamma_{122}^{(\alpha)} = 0.$$

From $\Gamma_{ij}^{k(\alpha)} = \Gamma_{ijm}^{(\alpha)} g^{km}$

$$\Gamma_{11}^{2(\alpha)} = \Gamma_{11m}^{(\alpha)} g^{2m} = \frac{(r+1)\{6 + (1-\alpha)(r-1)\}}{2rv(r+5)}$$

(2.11) $$\Gamma_{21}^{1(\alpha)} = \Gamma_{21m}^{(\alpha)} g^{1m} = \frac{-2(r+2) + (1-\alpha)(r-1)}{v(r+5)}$$

$$\Gamma_{21}^{2(\alpha)} = \Gamma_{11}^{1(\alpha)} = 0.$$

Since the α-curvature tensors $R_{ijkm}^{(\alpha)}$ are defined by

$$R_{ijkm}^{(\alpha)} = (\Gamma_{ik,j}^{s(\alpha)} - \Gamma_{jk,i}^{s(\alpha)})g_{sm} + \Gamma_{jrm}^{(\alpha)}\Gamma_{ik}^{r(\alpha)} - \Gamma_{irm}^{(\alpha)}\Gamma_{jk}^{r(\alpha)}$$

where $\Gamma_{ij}^{k(\alpha)} = \Gamma_{ijm}^{(\alpha)} g^{km}$,
by (2.3), (2.10) and (2.11)

$$R_{1212}^{(\alpha)} = (\partial_2 \Gamma_{11}^{s(\alpha)} - \partial_1 \Gamma_{21}^{s(\alpha)})g_{s2} + \Gamma_{2r2}^{(\alpha)}\Gamma_{11}^{r(\alpha)} - \Gamma_{1r2}^{(\alpha)}\Gamma_{21}^{r(\alpha)}$$

$$= (\partial_2 \Gamma_{11}^{2(\alpha)})g_{22} - \Gamma_{112}^{(\alpha)}\Gamma_{21}^{1(\alpha)} + \Gamma_{222}^{(\alpha)}\Gamma_{11}^{2(\alpha)}$$

$$= -\frac{(r+1)\{12(r+2) + 2(1-\alpha)(r-1)^2 - (1-\alpha)^2(r-1)^2\}}{v^4(r+3)(r+5)^2}$$

$\square$

Since the α-scalar curvature $K^{(\alpha)}$ is defined by

$$K^{(\alpha)} = \frac{1}{n(n-1)} R_{ijkm}^{(\alpha)} g^{im} g^{jk}, \quad K^{(\alpha)} = \frac{1}{2}(-2R_{1212}^{(\alpha)} g^{11} g^{22}).$$

Thus we have

Theorem 2.2. *The α-scalar curvature of the t-distribution is given by*

$$K^{(\alpha)} = \frac{(r+3)\{12(r+2) + 2(1-\alpha)(r-1)^2 - (1-\alpha)^2(r-1)^2\}}{2r(r+5)^2}.$$

Thus if $\alpha = 0$, we have the following corollary.

Corollary 2.3. *The scalar curvature R and the Gaussian curvature K of the t-distribution, are*

$$R = -\frac{r+3}{r}, \ K = \frac{1}{2}R.$$

3. Conclusions

The Fisher information matrix(FIM) measures the curvature of the log-likelihood surface. Flat surfaces around the maximum do not inspire high confidence in estimated parameter values, while steep surfaces lead to sharp estimates. It is important to know the shape of a statistical model in the whole set of probability distributions. The information content is large if the FIM is large, because the likelihood is sharply peaked. we are sure that the maximum likelihood(ML) solution is a good estimate. If the curvature is small, then the likelihood probability distribution is very broad. So the ML estimate is not as good because the variance is very large.

A one-parameter family of Affine connections are called the α-connections. The duality between the α-connection and the α-connection together with the metric play an essential role in this geometry. This kind of duality, having emerged from manifolds of probability distributions, is ubiquitous, appearing in a variety of problems which might have no explicit relation to probability theory. The notion of α-curvature serves an important role in the asymptotic theory of statistical estimation, ancillary statistics, conditional inference and Bartlett adjustment in the likelihood ratio test.

References

[1] Adbel-All, N. H., Abd-Ellah, H. N. and Moustafa, H. M. (2003). Information geometry and statistical manifold. *Chaos, Solitons and Fractals*, **15**, 161-172.

[2] Amari, S. (1982). Differential geometry of curved exponential families-curvatures and information loss. *Ann. Statist.* 10.

[3] Amari, S. (1985). Differential geometrical methods in statistics, *Springer Lecture Notes in Statistics*, **28**.

[4] Arwini, K. A. and Dodson, C. T. J. (2007). Alpha-geometry of the Weibull manifold. ***Second Basic Sciences Conference***, Al-Fatah University, Tripoli, Libya 4-8.

[5] Cho, B. S. and Baek, H. Y. (2006). Geometric properties of t-distribution. *Honam Mathematical Journal*. **28**, 433-438.

[6] Efron, B. (1975). Defining the curvature of a statistical problem. *Annual. Statistics*. **3**, 1109-1242.

[7] Kass, R. E. and Vos, P. W. (1997). Geometrical foundations of asymptotic inference, *John Wiley and Sons, Inc.*

[8] Kass, R. E. (1989). The geometry of asymptotic inference, *Statistical Science*, **4**, 188- 219.

[9] Murray, M. K. and Rice, J. W. (1993). Differential geometry and Statistics, *Chapman and Hall, New York.*

[10] Rao, C. R. (1945). Information and the accuracy attainable in estimation of statistical parameters, *Bull. Calcutta Math. Soc*, **37**, 81-91.

Bong Sik Cho
Division of Mathematics and Informational Statistics, Research institute
for basic sciences, Wonkwang University,
Iksan 570-749, Korea.
E-mail: bscho@wonkwang.ac.kr

Sun Young Jung
College of Engineering, Hanyang University,
Seoul 133-791, Korea.
E-mail: jsycbs@hanyang.ac.kr