美作女子大学・美作女子大学短期大学部紀要
1997, Vol. 42, 42～46

論　文

A Generalization of Hougaard-Kass's Differential Equation

Hougaard-Kass の微分方程式の一般化

Tsukio Morita

Abstract

For a one-dimensional curved exponential family, the differential equation derived by Hougaard(1982) gives four parametrizations which have the following properties: (1) normal likelihod, (2) variance stability, (3) zero asymptotic skewness of the MLE and (4) zero asymptotic bias of the MLE. Kass(1984) showed that Hougaard's differential equation can be expressed by the α-connection. This note generalizes Hougaard-Kass's results to a multi-parameter curved exponential family and gives conditions for the existence of solutions in the resulting partial differential equations.

INTRODUCTION

Let F be an m-dimensional curved exponential family

$$F = \{f(x;\theta(\beta)) = \exp\{t_i(x)\theta^i(\beta) - \chi(\theta(\beta))\}, \quad \beta \in B\},$$

where the parameter space B is an open subset of R^m and the mapping $\theta : B \to R^n$ is twice differentiable in β and χ is thrice differentiable in θ. We use Einstein's summation convention for the index i, i.e $t_i(x)\theta^i(\beta)$ stands for $\sum_{i=1}^n t_i(x)\theta^i(\beta)$.
In the one-parameter case Hougaard(1982) derived the differential equation

$$\frac{d^2\psi/d\beta^2}{d\psi/d\beta} = \left(\frac{1-\alpha}{2}\frac{\partial^3\chi}{\partial\theta^i\partial\theta^j\partial\theta^k}\frac{d\theta^i}{d\beta}\frac{d\theta^j}{d\beta}\frac{d\theta^k}{d\beta} + \frac{\partial^2\chi}{\partial\theta^i\partial\theta^j}\frac{d^2\theta^i}{d\beta^2}\frac{d\theta^j}{d\beta}\right) \Big/ \frac{\partial^2\chi}{\partial\theta^i\partial\theta^j}\frac{d\theta^i}{d\beta}\frac{d\theta^j}{d\beta}, \tag{1}$$

where $\psi : B \to R$ is twice continuously differentiable and $\frac{d\psi}{d\beta} \neq 0$. Then he elucidated that ψ has normal likelihood, variance stabilizing parameter, zero asymptotic skewness of the MLE and zero asymptotic bias of the MLE according as $\alpha = \frac{1}{3}, 0, -\frac{1}{3}$ and -1, respectively.

Kass(1984) showed that, putting the numerator of the right-hand side of (1) by $\overset{\alpha}{\Gamma}_\beta$, the differential equation (1) is equivalent to $\overset{\alpha}{\Gamma}_\psi = 0$.

In this note we aim to give partial differential equations which generalizes (1) to the case when β is an m-dimensional vector. Furthermore, we show some conditions for the existence of solutions in the new equations.

GENERALIZATION OF HOUGAARD KASS'S DIFFERETIAL EQUATION

We define the information metric and the α-connection in F as follows (see Amari, 1982a):

$$g_{ab}(\beta) = E\Big(\frac{\partial l}{\partial\beta^a}\frac{\partial l}{\partial\beta_b}\Big)$$

and

$$\overset{\alpha}{\Gamma}_{abc}(\beta) = E\Big(\frac{\partial^2 l}{\partial\beta^a\partial\beta^b}\frac{\partial l}{\partial\beta^c}\Big) + \frac{1-\alpha}{2}E\Big(\frac{\partial l}{\partial\beta^a}\frac{\partial l}{\partial\beta^b}\frac{\partial l}{\partial\beta^c}\Big),$$

where $l(\beta) = \log f(x;\theta(\beta))$ and the indicies $a, b, c, etc.$ move from 1 to m.

Theorem 1. *Suppose the parameter space B is an open subset of R^m and the mapping $\psi : B \to R^m$ is C^2-diffeomorphism. Consider the partial differential equations*

$$\frac{\partial\psi^d}{\partial\beta^e}\Big(\overset{1/3}{\Gamma}_{abc}(\beta) + \overset{1/3}{\Gamma}_{bca}(\beta) + \overset{1/3}{\Gamma}_{abc}(\beta)\Big) = \frac{\partial^2\psi^d}{\partial\beta^a\partial\beta^b}g_{ec} + \frac{\partial^2\psi^d}{\partial\beta^b\partial\beta^c}g_{ea} + \frac{\partial^2\psi^d}{\partial\beta^c\partial\beta^a}g_{eb} \quad (2)$$

and

$$\frac{\partial\psi^d}{\partial\beta^e}\overset{\alpha}{\Gamma}_{abc}(\beta) = \frac{\partial^2\psi^d}{\partial\beta^a\partial\beta^b}g_{ec} \quad (\alpha = 0, -\frac{1}{3} \text{ or } -1). \quad (3)$$

(a)*ψ has a normal likelihood if and only if ψ is a solution to* (2), (b) *ψ is a variance stbilizing parameter if and only if ψ is a solution to* (3) *with $\alpha = 0$,* (c) *ψ makes the asymptotic skewness of the MLE to be zero if and only if ψ is a solution to* (3) *with $\alpha = -\frac{1}{3}$ and* (d) *ψ makes the asymptotic bias of the MLE to be zero if ψ is a solution to* (3) *with $\alpha = -1$.*

Remark 1. When $f(x;\theta(\beta))$ is a density in a one -dimensional curved exponential family, Theorem 1 coincides with Hougaard's results. In fact the partial differential equations (2) and (3) reduce to the differential equation

$$\frac{d^2\psi/d\beta^2}{d\psi/d\beta} = \overset{\alpha}{\Gamma}_{111}(\beta)g^{11}, \quad (4)$$

where $\psi = \psi^1, \beta = \beta^1$

$$\overset{\alpha}{\Gamma}_{111}(\beta) = \frac{1-\alpha}{2}\frac{\partial^3\chi}{\partial\theta^i\partial\theta^j\partial\theta^k}\frac{d\theta^i}{d\beta}\frac{d\theta^j}{d\beta}\frac{d\theta^k}{d\beta} + \frac{\partial^2\chi}{\partial\theta^i\partial\theta^j}\frac{d^2\theta^i}{d\beta^2}\frac{d\theta^j}{d\beta}, \quad (5)$$

and

$$g^{11} = \frac{1}{\frac{\partial^2\chi}{\partial\theta^i\partial\theta^j}\frac{d\theta^i}{d\beta}\frac{d\theta^j}{d\beta}}. \quad (6)$$

Substituting the values of $\overset{\alpha}{\Gamma}_{111}$ and g^{11} given by (5) and (6) into (4), we have Hougaard's differential equation.

Remark 2. Note that the partial differential equations (2) and (3) with $\alpha = 0$ hold for any regular parametric families.

Before we give the proof of Theorem 1, we explain the variance stabilizing parameter ψ. Suppose $\sqrt{N}(T_N - \beta)$ has an asymptotic mutivariate normal distribution with zero mean vector and a variance matrix (g^{ab}), then by a one-to-one mapping ψ, $\sqrt{N}(\psi(T_N) - \psi(\beta))$ has an asymptotic multivariate normal distribution with zero mean vector and a variance matrix (g'^{ab}), where $g'^{ab} = \frac{\partial\psi^a}{\partial\beta^c}\frac{\partial\psi^b}{\partial\beta^d}g^{cd}$ and (g^{cd}) is the inverse matrix of (g_{cd}). When (g'^{ab}) is a constant matrix we call ψ the variance stabilizing parameter.

Proof of Theorem 1

(a) Normal likelihood

Partially differentiating the both sides of the equation $E\left(\frac{\partial l}{\partial\beta^a}\right) = 0$ by β^b andβ^c, we get

$$E\left(\frac{\partial^3 l}{\partial\beta^a\partial\beta^b\partial\beta^c} + \frac{\partial^2 l}{\partial\beta^a\partial\beta^b}\frac{\partial l}{\partial\beta^c} + \frac{\partial^2 l}{\partial\beta^b\partial\beta^c}\frac{\partial l}{\partial\beta^a} + \frac{\partial^2 l}{\partial\beta^c\partial\beta^a}\frac{\partial l}{\partial\beta^b} + \frac{\partial l}{\partial\beta^a}\frac{\partial l}{\partial l\beta^b}\frac{\partial l}{\partial\beta^c}\right) = 0 \quad (7)$$

and hence

$$E\left(\frac{\partial^3 l}{\partial\beta^a\partial\beta^b\partial\beta^c}\right) = -E\Bigg(\left(\frac{\partial^2 l}{\partial\beta^a\partial\beta^b}\frac{\partial l}{\partial\beta^c} + \frac{1}{3}\frac{\partial l}{\partial\beta^a}\frac{\partial l}{\partial l\beta^b}\frac{\partial l}{\partial\beta^c}\right)$$
$$+\left(\frac{\partial^2 l}{\partial\beta^b\partial\beta^c}\frac{\partial l}{\partial\beta^a} + \frac{1}{3}\frac{\partial l}{\partial\beta^a}\frac{\partial l}{\partial l\beta^b}\frac{\partial l}{\partial\beta^c}\right)$$
$$+\left(\frac{\partial^2 l}{\partial\beta^b\partial\beta^c}\frac{\partial l}{\partial\beta^a} + \frac{1}{3}\frac{\partial l}{\partial\beta^a}\frac{\partial l}{\partial l\beta^b}\frac{\partial l}{\partial\beta^c}\right)\Bigg)$$
$$= -\left(\overset{1/3}{\Gamma}_{abc}(\beta) + \overset{1/3}{\Gamma}_{bca}(\beta) + \overset{1/3}{\Gamma}_{abc}(\beta)\right).$$

That ψ has a normal likelihood, i.e. $E\left(\frac{\partial^3 l}{\partial\psi^a\partial\psi^b\partial\psi^c}\right) = 0$ for all ψ,is equivalent to

$$\overset{1/3}{\Gamma'}_{abc} + \overset{1/3}{\Gamma'}_{bca} + \overset{1/3}{\Gamma'}_{abc} = 0 \quad for\ all\ \psi, \tag{8}$$

where $\overset{\alpha}{\Gamma'}_{abc}$ is the α-connection for ψ. Using the transformation-formula of the α-connection, we obtain

$$\overset{1/3}{\Gamma'}_{abc} = \frac{\partial\beta^{a'}}{\partial\psi^a}\frac{\partial\beta^{c'}}{\partial\psi^b}\left(\frac{\partial\beta^{t'}}{\partial\psi^c}\overset{1/3}{\Gamma}_{a'c't'} - \frac{\partial\beta^{e'}}{\partial\psi^e}\frac{\partial\beta^{t'}}{\partial\psi^c}\frac{\partial^2\psi^e}{\partial\beta^{a'}\partial\beta^{t'}}g_{e't'}\right),$$
$$\overset{1/3}{\Gamma'}_{bca} = \frac{\partial\beta^{a'}}{\partial\psi^b}\frac{\partial\beta^{c'}}{\partial\psi^c}\left(\frac{\partial\beta^{t'}}{\partial\psi^a}\overset{1/3}{\Gamma}_{a'c't'} - \frac{\partial\beta^{e'}}{\partial\psi^e}\frac{\partial\beta^{t'}}{\partial\psi^a}\frac{\partial^2\psi^e}{\partial\beta^{a'}\partial\beta^{t'}}g_{e't'}\right), \tag{9}$$
$$\overset{1/3}{\Gamma'}_{cab} = \frac{\partial\beta^{a'}}{\partial\psi^c}\frac{\partial\beta^{c'}}{\partial\psi^a}\left(\frac{\partial\beta^{t'}}{\partial\psi^b}\overset{1/3}{\Gamma}_{a'c't'} - \frac{\partial\beta^{e'}}{\partial\psi^e}\frac{\partial\beta^{t'}}{\partial\psi^b}\frac{\partial^2\psi^e}{\partial\beta^{a'}\partial\beta^{t'}}g_{e't'}\right).$$

Substitution of (9) into (8) leads us to (2), which completes the proof.
(b) Variance stability, zero asymptotic skewness and zero asymptotic bias
Note that (g'^{ab}) is a constant matrix if and only if $\frac{\partial g'^{ab}}{\partial\psi^c} = 0$ and it is equivalent to $\frac{\partial g'_{ab}}{\partial\psi^c} = 0$. Now we have

$$\frac{\partial g'_{ab}}{\partial\psi^c} = \frac{\partial}{\partial\psi^c}E\left(\frac{\partial l}{\partial\psi^a}\frac{\partial l}{\partial\psi^b}\right) = E\Bigg(\left(\frac{\partial^2 l}{\partial\psi^a\partial\psi^c}\frac{\partial l}{\partial\psi^b} + \frac{1}{2}\frac{\partial l}{\partial\psi^a}\frac{\partial l}{\partial\psi^b}\frac{\partial l}{\partial\psi^c}\right)$$
$$+\left(\frac{\partial^2 l}{\partial\psi^c\partial\psi^b}\frac{\partial l}{\partial\psi^a} + \frac{1}{2}\frac{\partial l}{\partial\psi^a}\frac{\partial l}{\partial\psi^b}\frac{\partial l}{\partial\psi^c}\right)\Bigg)$$
$$= \overset{0}{\Gamma'}_{acb}(\psi) + \overset{0}{\Gamma'}_{cba}(\psi).$$

From

$$\overset{0}{\Gamma'}_{abc} = \frac{1}{2}\left(\frac{\partial g'_{bc}}{\partial\psi^a} + \frac{\partial g'_{ac}}{\partial\psi^b} - \frac{\partial g'_{ab}}{\partial\psi^c}\right),$$

we find that $\overset{0}{\Gamma'}_{acb} + \overset{0}{\Gamma'}_{cba} = 0$ is equivalent to $\overset{0}{\Gamma'}_{abc} = 0$. Thus a necessary and sufficient condition for the existence of the variance stabilizing parameter is $\overset{0}{\Gamma'}_{abc} = 0$.

Similary as the proof of (a), we can easily obtain (3) with $\alpha = 0$. The equation with $\alpha = -\frac{1}{3}$ or 1 in (3) hold for the MLE for a curved exponential family. Amari showed that

$$K_{abc} = -3 \overset{-1/3}{\Gamma}{}_{abc}$$

and

$$b^a = -\frac{1}{2} \overset{-1}{\Gamma}{}^a_{bc}\, g^{bc},$$

where K_{abc} and b^a are, respectively, an asymptotic skewness and bias (see Amari, 1982a(6.2) and b(4.2) for details). A use of these results and the transformation of the α-connection derives the desired results.

CODITIONS FOR THE EXISTENCE OF SOLUTIONS

We shall find conditions for the existence of respective parametrizations satisfying (1) normal likelihood, (2) variance stability, (3) zero asymptotic skewness of the MLE and (4) zero aymptotic bias of the MLE. It is easily seen that for (1) and (4) sufficient conditions for the existence of solutions are $\overset{1/3}{\Gamma'}{}_{abc}= 0$ and $\overset{-1}{\Gamma'}{}_{abc}= 0$ and that for (2) and (3) $\overset{0}{\Gamma'}{}_{abc}= 0$ and $\overset{-1/3}{\Gamma'}{}_{abc}= 0$ give necessary and sufficient conditions. $\overset{\alpha}{\Gamma'}{}_{abc}= 0$ $(\alpha = \frac{1}{3}, 0, -\frac{1}{3}\ or\ -1)$ is equivalent to the partial differential equation

$$\frac{\partial \psi^e}{\partial \beta^b} \overset{\alpha}{\Gamma'}{}^b_{ac} = \frac{\partial^2 \psi^e}{\partial \beta^a \partial \beta^c}, \tag{10}$$

where $\overset{\alpha}{\Gamma'}{}^b_{ac} = \overset{\alpha}{\Gamma'}{}_{acd}\, g^{bd}$. Putting $\psi^e_b = \frac{\partial \psi^e}{\partial \beta^b}$, (10) becomes equivalent to

$$\psi^e_b = \frac{\partial \psi^e}{\partial \beta^b}$$

$$\frac{\partial \psi^e_b}{\partial \beta^c} = \psi^e_l \overset{\alpha}{\Gamma'}{}^l_{bc}\,. \tag{11}$$

Now we have

$$\frac{\partial}{\partial \beta^a}\left(\frac{\partial \psi^e}{\partial \beta^b}\right) - \frac{\partial}{\partial \beta^b}\left(\frac{\partial \psi^e}{\partial \beta^a}\right) = \frac{\partial \psi^e_b}{\partial \beta^a} - \frac{\partial \psi^e_a}{\partial \beta^b} = \psi^e_d\left(\overset{\alpha}{\Gamma}{}^d_{ba} - \overset{\alpha}{\Gamma}{}^d_{ab}\right) = 0,$$

since $\overset{\alpha}{\Gamma}{}^d_{ab} = \overset{\alpha}{\Gamma}{}^d_{ba}$. Furthermore,

$$\frac{\partial}{\partial \beta^f}\left(\frac{\partial \psi^e_a}{\partial \beta^h}\right) - \frac{\partial}{\partial \beta^h}\left(\frac{\partial \psi^e_a}{\partial \beta^f}\right) = \frac{\partial}{\partial \beta^f}\left(\psi^e_l \overset{\alpha}{\Gamma}{}^l_{ah}\right) - \frac{\partial}{\partial \beta^h}\left(\psi^e_l \overset{\alpha}{\Gamma}{}^l_{af}\right)$$

$$= \psi^e_d\left(\frac{\partial}{\partial \beta^f}\overset{\alpha}{\Gamma}{}^d_{ah} - \frac{\partial}{\partial \beta^h}\overset{\alpha}{\Gamma}{}^d_{af} + \overset{\alpha}{\Gamma}{}^d_{lf}\overset{\alpha}{\Gamma}{}^l_{ah} - \overset{\alpha}{\Gamma}{}^d_{lh}\overset{\alpha}{\Gamma}{}^l_{af}\right)$$

$$= \psi^e_d \overset{\alpha}{R}{}^d_{fha},$$

where

$$\overset{\alpha}{R}{}^{d}_{fha}= \frac{\partial}{\partial\beta^f}\overset{\alpha}{\Gamma}{}^{d}_{ah} - \frac{\partial}{\partial\beta^h}\overset{\alpha}{\Gamma}{}^{d}_{af} + \overset{\alpha}{\Gamma}{}^{d}_{lf}\overset{\alpha}{\Gamma}{}^{l}_{ah} - \overset{\alpha}{\Gamma}{}^{d}_{lh}\overset{\alpha}{\Gamma}{}^{l}_{af}$$

is called an α-Riemann-Christofel curvature.

If $\overset{\alpha}{R}{}^{d}_{fha}= 0$, the equation (11) is completely integrable. Hence there exist a solution in the neighborfood of any initial value. Therefore we can obtain the following theorem.

Theorem 2. (a) *A sufficient condition for the existence of solutions of the equations* (2) *and* (3) *with* $\alpha = -1$ *is that*

$$\overset{\alpha}{R}{}^{d}_{abc}= 0 \quad (\alpha = \frac{1}{3}\ or\ -1) \tag{12}$$

(b)*A necessary and sufficient condition for the existence of the solutions of the equations* (3)*with* $\alpha = 0$ *or* $-\frac{1}{3}$ *is that*

$$\overset{\alpha}{R}{}^{d}_{abc}= 0 \quad (\alpha = -\frac{1}{3}\ or\ 0)$$

Corollary. *A one-parameter exponential family is completely integrable.*
This statement is obvious, since

$$\overset{\alpha}{R}{}^{1}_{111}= \frac{\partial}{\partial\beta^1}\overset{\alpha}{\Gamma}{}^{1}_{11} - \frac{\partial}{\partial\beta^1}\overset{\alpha}{\Gamma}{}^{1}_{11} + \overset{\alpha}{\Gamma}{}^{1}_{11}\overset{\alpha}{\Gamma}{}^{1}_{11} - \overset{\alpha}{\Gamma}{}^{1}_{11}\overset{\alpha}{\Gamma}{}^{1}_{11}= 0.$$

Remark 3. In this case it is easily seen that (12) becomes a necessary and sufficient condition.

References

[1] Amari, S. (1982a). Geometrical theory of asymptotic ancillarity and conditional inference. *Biometrika* , **69**, 1–7

[2] Amari, S. (1982b). Differential geometry of estimation: higher-order efficiency in curved exponential families. *Technical reports, Tokyo Univ.*

[3] Hougaard, P. (1982). Parametrizations of non-linear models. *J. R. Statist. Soc. B*, **44**, 244–252

[4] Kass, R. (1984). Canonical parametrizations and zero parameter-effects cuevature. *J. R. Statist. B*, **46**, 86–92

（1996年12月2日　受理）