On a characterization of a class of functions defined on the space of positive definite matrices

Dedicated to Professor Shigeru Furuya on his 60th birthday

By Nagayoshi IWAHORI

Introduction.

It is pointed out in Kimura [1] that a test-function defined on the set of variance-covariance matrices is of use only when the result of comparisons by the function values are independent of the observation-coordinate-system. In [1] it is claimed that such a function should be a function of the determinant of the matrix under several strong conditions on the function. The content of [1] were reported in a conference at Hiroshima University, 1971. Attending Kimura's talk at this conference, I have given several remarks to clarify the formulation of the main statement and to strengthen the result as was given in this note. I wish to record here the content of my short reporting talk in that conference.

In our formulation, the problem is to characterize the equivalence class of the determinant function Δ defined on the space X_n of all positive definite symmetric matrices of degree n, in the sense of the following equivalence $\sim$ of functions $f: X_n \to \boldsymbol{R}$, $g: X_n \to \boldsymbol{R}$, namely, $f \sim g$ means that for $x, y \in X$, (i) $f(x) > f(y)$ if and only if $g(x) > g(y)$ and (ii) $f(x) = f(y)$ if and only if $g(x) = g(y)$.

Our main result is that the class of Δ is characterized by the following two properties: (1) The class is stable under the action of $GL(n, \boldsymbol{R})$: the action being naturally induced by that of $GL(n, \boldsymbol{R})$ on X_n: $a \to \sigma a\, {}^t\sigma$, $a \in X_n$, $\sigma \in GL(n, \boldsymbol{R})$, (2) no function f in the class attains its minimum on X_n; but f satisfies

$$\lim_{x \to 0} f(x) = \inf_{x \in X_n} f(x) .$$

§ 1.

Let X be a Hausdorff space. We denote by $\boldsymbol{R}^X$ the set of all real-valued functions defined on X. Let $f \in \boldsymbol{R}^X$, $g \in \boldsymbol{R}^X$. We say that f is equivalent to g (in notation $f \sim g$), if the following conditions (i), (ii) are satisfied:

(i) for any $x, y \in X$, $f(x) = f(y)$ if and only if $g(x) = g(y)$

(ii) for any $x, y \in X$, $f(x) > f(y)$ if and only if $g(x) > g(y)$.

It is then easy to see that $\sim$ is an equivalence relation. We note that $f \sim \exp(f)$. Thus every equivalence class contains a function g such that $g(x) > 0$ for every

$x \in X$.

Now let X_n be the space of all positive definite real symmetric matrices of degree n. Let us denote by Δ the determinant function on X: $\Delta(a)=\det(a)$, $a \in X_n$. The general linear group $G_n = GL(n, \boldsymbol{R})$ acts on X in the usual manner: for $\sigma \in G_n$, $x \in X_n$, $x^\sigma = {}^t\sigma \cdot x \cdot \sigma$, ${}^t\sigma$ being the transposed of σ. Thus G acts also on $\boldsymbol{R}^{X_n}$ as follows: for $\sigma \in G_n$, $f \in \boldsymbol{R}^{X_n}$, $f_\sigma(x) = f(x^\sigma)$.

We denote by $M_n(\boldsymbol{R})$ the linear space of all real matrices of degree n. Then $X_n \subset M_n(\boldsymbol{R})$ and the zero matrix 0 belongs to the closure $\bar{X}_n$ of X_n in $M_n(\boldsymbol{R})$. Hence we can talk about the existence of

$$\lim_{x \to 0} f(x)$$

for $f \in \boldsymbol{R}^{X_n}$.

THEOREM. *Let* $f \in \boldsymbol{R}^{X_n}$. *Suppose that* $f(x) > 0$ *for every* $x \in X_n$. *Then* $f \sim \Delta$ *if and only if the following conditions* I, II *are satisfied.*

I. $f \sim f_\sigma$ *for any* $\sigma \in GL(n, \boldsymbol{R})$.

II. *Let* $m = \inf_{x \in X_n} f(x)$. *Then* $f(x) > m$ *for every* $x \in X_n$. *Furthermore* $\lim_{x \to 0} f(x)$ *exists and is equal to* m.

PROOF. *Necessity of* I, II. Suppose $f \sim \Delta$. Then $f(x) > f(y)$ implies $\Delta(x) > \Delta(y)$. Hence $\Delta({}^t\sigma x \sigma) > \Delta({}^t\sigma y \sigma)$ for every $\sigma \in G_n$. Then $f({}^t\sigma x \sigma) > f({}^t\sigma y \sigma)$ i.e. $f_\sigma(x) > f_\sigma(y)$. Similarly $f(x) = f(y)$ implies $f_\sigma(x) = f_\sigma(y)$. Thus we get $f \sim f_\sigma$. Suppose now that there exists a point $x \in X_n$ such that $f(x) = m$. Then $\Delta\left(\frac{1}{2}x\right) < \Delta(x)$ implies $m \leqq f\left(\frac{1}{2}x\right) < f(x) = m$, which is impossible. Hence we have $f(x) > m$ for every $x \in X_n$. Finally assume that $\lim_{x \to 0} f(x) = m$ does not hold. Then there exists a positive number ε_0 and a sequence $x_1, x_2, \cdots$ of points in X_n such that $\lim x_\nu = 0$ and $f(x_\nu) \geqq m + \varepsilon_0$ $(\nu = 1, 2, \cdots)$. Take a point $a \in X_n$ which satisfies $f(a) < m + \varepsilon_0$. Now $\lim x_\nu = 0$ implies $\lim \Delta(x_\nu) = 0$. Hence $\Delta(x_k) < \Delta(a)$ for sufficiently large k: therefore $f(x_k) < f(a) < m + \varepsilon_0$ which is impossible.

Sufficiency of I, II. Suppose now f satisfies I and II. We begin with the following

LEMMA 1. *Define a subgroup* H *of* $GL(n, \boldsymbol{R})$ *by*

$$H = \{\sigma \in GL(n, \boldsymbol{R}) \mid f = f_\sigma\}.$$

Then every element of finite order in $GL(n, \boldsymbol{R})$ *is contained in* H.

PROOF. Suppose $\sigma \in GL(n, \boldsymbol{R})$, $\sigma \notin H$. Then $f \neq f_\sigma$ implies the existence of a point $x_0 \in X$ such that $f(x_0) \neq f_\sigma(x_0)$. Suppose for example that $f(x_0) < f_\sigma(x_0) = f({}^t\sigma x_0 \sigma)$. Then $f \sim f_\sigma$ implies $f_\sigma(x_0) < f_\sigma({}^t\sigma x_0 \sigma)$, i.e. $f({}^t\sigma x_0 \sigma) < f({}^t\sigma^2 x_0 \sigma^2)$, which in turn implies $f_\sigma({}^t\sigma x_0\sigma)<f_\sigma({}^t\sigma^2 x_0\sigma^2)$, i.e. $f({}^t\sigma^2 x_0\sigma^2)<f({}^t\sigma^3 x_0\sigma^3)$, and so on. Thus we get an infinite sequence

$$f(x_0)<f({}^t\sigma x_0\sigma)<f({}^t\sigma^2 x_0\sigma^2)<\cdots.$$

However this implies that the order of σ can not be finite, q.e.d.

Now let us denote by F the subgroup of $GL(n, \boldsymbol{R})$ generated by all elements of finite order in $GL(n, \boldsymbol{R})$. Then F is a normal subgroup of $GL(n, \boldsymbol{R})$; furthermore $F\subset H$ by Lemma 1. $F\cap SL(n, \boldsymbol{R})$ is a normal subgroup of $SL(n, \boldsymbol{R})$. Now let us quote the following well-known classical result:

LEMMA 2. *Suppose* $n\geqq 2$. *Then the normal subgroups of* $SL(n, \boldsymbol{R})$ *are*

$$SL(n, \boldsymbol{R}),\ \{1\},\ \{1,-1\}.$$

($\{1,-1\}$ *is the case only when* n *is even.*)

See, e.g. [2] for the proof.

Now the subgroup $F\cap SL(n, \boldsymbol{R})$ can not coincide with $\{1\}$ or with $\{1,-1\}$ if $n>1$, since

$$\left(\begin{array}{cc|ccc} 0 & 1 & & 0 & \\ -1 & 0 & & & \\ \hline & & 1 & & 0 \\ & 0 & & \ddots & \\ & & 0 & & 1 \end{array}\right)$$

is in $F\cap SL(n, \boldsymbol{R})$. Hence $F\cap SL(n, \boldsymbol{R})=SL(n, \boldsymbol{R})$, i.e. $F\supset SL(n, \boldsymbol{R})$. This is also true for $n=1$.

Now consider the subgroup

$$SL^{\pm}(n, \boldsymbol{R})=\{\sigma\in GL(n, \boldsymbol{R})\,|\,\det(\sigma)=\pm 1\}$$

of $GL(n, \boldsymbol{R})$. Since $SL^{\pm}(n, \boldsymbol{R})$ is generated by $SL(n, \boldsymbol{R})$ and an element

$$\left(\begin{array}{ccc|c} 1 & & & \\ & \ddots & & 0 \\ & & 1 & \\ \hline & 0 & & -1 \end{array}\right)$$

in F, we have $SL^{\pm}(n, \boldsymbol{R})\subset F$.

Let now $a\in GL(n, \boldsymbol{R})$ be of finite order. Then $\Delta(a)$ is real and is a root of unity. Hence $\Delta(a)=\pm 1$. Hence $a\in SL^{\pm}(n, \boldsymbol{R})$. Therefore $F\subset SL^{\pm}(n, \boldsymbol{R})$. Thus we have shown that $F=SL^{\pm}(n, \boldsymbol{R})$. Since $F\subset H$ by Lemma 1, we have proved the following.

LEMMA 3. $$SL^{\pm}(n, \boldsymbol{R})\subset H.$$

We now proceed to show $f\sim\Delta$.

LEMMA 4. *Let* $x_0\in X_n$, $y_0\in X_n$. *Suppose* $\Delta(x_0)=\Delta(y_0)$. *Then* $f(x_0)=f(y_0)$.

PROOF. As is well-known, $GL(n, \boldsymbol{R})$ acts transitively on X_n. So there exists an element $\sigma_0\in GL(n, \boldsymbol{R})$ such that ${}^t\sigma_0 x_0\sigma_0=y_0$. Taking determinants of both sides, we see $\Delta(\sigma_0)=\pm 1$, using $\Delta(x_0)=\Delta(y_0)$. Hence $\sigma_0\in SL^{\pm}(n, \boldsymbol{R})=F\subset H$, i.e. $f=f_{\sigma_0}$. Hence $f({}^t\sigma_0 x_0\sigma_0)=f(x_0)$, i.e. $f(x_0)=f(y_0)$, q.e.d.

LEMMA 5. *Let* $x_0\in X_n$, $y_0\in X_n$. *Suppose* $\Delta(x_0)>\Delta(y_0)$. *Then* $f(x_0)>f(y_0)$.

PROOF. Take an element $\sigma_0\in GL(n, \boldsymbol{R})$ such that ${}^t\sigma_0 x_0\sigma_0=y_0$. Then $\Delta(x_0)>\Delta(y_0)$ implies $\Delta(\sigma_0)^2<1$. Choose $\rho\in\boldsymbol{R}$ such that $0<\rho<1$, $\Delta(\sigma_0)^2=\rho^2$ and put $\tau=\rho^{2/n}$. Then $0<\tau<1$ and $\Delta(\tau x_0)=\tau^n\Delta(x_0)=\rho^2\Delta(x_0)=\Delta(y_0)$. Hence we get $f(\tau x_0)=f(y_0)$ by Lemma 4. Now suppose that $f(x_0)\leqq f(y_0)$. Then $f(x_0)\leqq f(\tau x_0)$. Put $\tau_0=\sqrt{\tau}\,I$. Then $\tau x_0={}^t\tau_0 x_0\tau_0$. Now $f(x_0)\leqq f(\tau x_0)$ and $f\sim f_{\tau_0}$ imply $f_{\tau_0}(x_0)\leqq f_{\tau_0}(\tau x_0)$, i.e. $f({}^t\tau_0 x_0\tau_0)\leqq f({}^t\tau_0\tau x_0\tau_0)$, i.e. $f(\tau x_0)\leqq f(\tau^2 x_0)$, which in turn implies $f(\tau^2 x_0)\leqq f(\tau^3 x_0)$ and so on. Thus we get an infinite sequence

$$f(x_0)\leqq f(\tau x_0)\leqq f(\tau^2 x_0)\leqq \cdots.$$

Since $\lim\tau^\nu x=0$, we must have $\lim_{\nu\to\infty} f(\tau^\nu x)=m=\inf_{x\in X_n} f(x)$ by the condition II. Hence we get $f(x_0)\leqq m$. However then we have $f(x_0)=m$, contrary to the validity of II, q.e.d.

The proof of Theorem is now complete by Lemmas 4 and 5.

Added in proof. The referee has pointed out that the equality $SL^{\pm}(n, \boldsymbol{R})=F$ can be shown without having recourse to [2]. Namely using the elementary facts that (i) $SL(n, \boldsymbol{R})$ is generated by the conjugates of the element

$$\left(\begin{array}{cc|cccc} 1 & 1 & & & & \\ 0 & 1 & & & 0 & \\ \hline & & 1 & & & \\ & & & \ddots & & 0 \\ & 0 & & & \ddots & \\ & & 0 & & & 1 \end{array}\right)$$

and (ii) $\begin{pmatrix}-1 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$, it is easy to show that $SL^{\pm}(n, \boldsymbol{R})\subset F$. The converse inclusion is obvious.

I agree with the referee's opinion and I express my gratitude to the referee. But I would like to keep the present proof, since it seems to me plausible that there exist analogous facts (as stated in the theorem) for each irreducible symmetric Riemannian manifold $M=G/K$ of noncompact type associated with a real simple Lie group G of normal type (i. e. of Chevalley type) and a maximal compact subgroup K of G. In establishing analogous facts for this case, probably one would have recourse to the simplicity of the factor group G/Z, where Z is the center of G.

References

[1] Kimura, Takeo, On the best observation-coordinate-system, J. Japan Statist. Soc., **2** (1971), 19-26 (in Japanese).
[2] Chevalley, C., Sur certain groupes simples, Tôhoku Math. J., **7** (1955), 14-66.

(Received December 13, 1976)

Department of Mathematics
Faculty of Science
University of Tokyo
Hongo, Tokyo
113 Japan