UNIVERSAL ESTIMATORS OF A VECTOR PARAMETER

by

A. L. RUKHIN
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #78-25

January 2, 1979

Universal Estimators of a Vector Parameter

by

A. L. Rukhin

ABSTRACT

Let $\underline{x}$ be a random sample with a distribution depending on a vector parameter $\theta \in \mathbb{R}^m$. The description of distributions and generalized prior densities on $\mathbb{R}^m$ is given, for which the generalized Bayes estimator of θ, based on $\underline{x}$, is the same for all symmetric loss functions. These distributions form a special subclass of exponential family. The specification of this result for the case of a location parameter is considered. The proof of the main theorem is based on the solution of a functional equation of D'Alembert's type.

UNIVERSAL ESTIMATORS OF A VECTOR PARAMETER*

A. L. Rukhin
Purdue University

1. Introduction

Let P_θ, $\theta \in \Theta$, be a family of probability measures on an abstract space X, such that each distribution P_θ is absolutely continuous with respect to some σ-finite measure μ on X. We assume throughout the paper that the coincidence of distributions P_{θ_1} and P_{θ_2} imples $\theta_1=\theta_2$, and that Θ is an open connected subset of the Euclidean space $\mathbb{R}^m$. Let λ be a generalized prior density on Θ, and define

$$\pi_{\underline{x}}(\theta) = (\prod_1^n p(x_j,\theta))\lambda(\theta)$$

where $p(x,\theta) = \frac{dP_\theta}{d\mu}(x)$, $x \in X$, $\underline{x} = (x_1,\ldots,x_n) \in X^n$, and $\theta \in \Theta$. For convenience, we let $\pi_{\underline{x}}(\theta)$ be zero for $\theta \notin \Theta$. Also, let $W(\delta,\theta) = W(\delta-\theta)$ be the loss function depending only on the difference between the estimator δ and the true value of the parameter θ. Thus $W(t)$ is defined for $t \in \Theta-\Theta$, and we assume that for each w the set $\{t: W(t) \leq w\}$ is convex.

The generalized Bayes estimator $\delta(\underline{x})$ of θ based on the random sample $\underline{x}$ satisfies the equation

$$\int_\Theta W(\delta(\underline{x})-\theta)\pi_{\underline{x}}(\theta)d\theta = \inf_{t\in\Theta} \int_\Theta W(t-\theta)\pi_{\underline{x}}(\theta)d\theta. \qquad (1.1)$$

In general this estimator depends on the choice of the loss function W, which is rarely known exactly to the statistician. Therefore, it seems rather natural to investigate situations in which the generalized Bayes estimator is the same for every loss from a certain set of loss functions under consideration.

AMS 1970 Subject classification: Primary 62C10, Secondary 62F10, 62H05.

Key words and phrases: generalized Bayes estimators, CS set of loss functions, universal estimators, exponential family, functional equation of the D'Alembert's type.

*Research supported by the National Science Foundation under Grant No. MCS 77-19640.

This problem for the case m = 1 was solved by the author (Rukhin 1978b). The situation where θ is a real location parameter and $\lambda(\theta)$ is constant has been considered in [5]. In this paper we treat the case of arbitrary m.

If $\mathcal{W}$ is a set of loss functions W such that the integrals in (1.1) converge for every $W \in \mathcal{W}$, then the estimator $\delta(\underline{x})$ is called universal if $\delta(\underline{x})$ satisfies (1.1) for all $W \in \mathcal{W}$. Thus a universal estimator is optimal with regard to every loss function W from $\mathcal{W}$.

Assume that all functions W under consideration are symmetric (i.e. $W(-t) = W(t)$) and differentiable. We also suppose that differentiation with respect to t in the right side of (1.1) is allowed under the integral sign, and that the relation

$$\int W_i'(t)g(t)dt = 0$$

($W_i'(t) = \frac{\partial}{\partial t_i} W(t)$, i=1,..,m) valid for all $W \in \mathcal{W}$ and a continuous function g(t), implies that $g(t) = g(-t)$. Any class $\mathcal{W}$ of loss functions satisfying these conditions we call a CS (complete symmetric) set.

If $\delta(\underline{x})$ is universal with regard to a CS set $\mathcal{W}$, and $\pi_{\underline{x}}(t)$ is a continuous function of t, then

$$\int W_i'(t-\delta(\underline{x}))\pi_{\underline{x}}(t)dt = 0.$$

Hence for all $t \in \mathbb{R}^m$

$$\pi_{\underline{x}}(\delta(\underline{x})-t) = \pi_{\underline{x}}(\delta(\underline{x})+t). \tag{1.2}$$

Also, if the second derivatives of W(t) exist, the matrix $(\int \frac{\partial^2}{\partial t_i \partial t_k} W(t)\pi_{\underline{x}}(\delta(\underline{x})+t)dt)$ is positive semidefinite. This establishes the first part of the following proposition.

Proposition 1.1 Assume that $\delta(\underline{x})$ is the universal estimator of $\theta \in \Theta \subset \mathbb{R}^m$ with regard to a CS set $\mathcal{W}$ of loss functions. If $\pi_{\underline{x}}(t)$ is a continuous function of $t \in \mathbb{R}^m$, then (1.2) holds for all $t \in \mathbb{R}^m$. If (1.2) is satisfied for almost all t, then $\delta(\underline{x})$ is the generalized Bayes estimator for every symmetric convex loss function such that the integrals in (1.1) exist.

The second part of Proposition 1.1 is well known (cf. for instance Deutsch (1965) pp. 14,16). Its proof consists in noticing that if $W(-t) = W(t)$ and (1.2) holds, then

$$\begin{aligned} &\int W(t-\theta)\pi_{\underline{x}}(\theta)d\theta - \int W(\delta(\underline{x})-\theta)\pi_{\underline{x}}(\theta)d\theta \\ &= \int W(\theta+t')\pi_{\underline{x}}(\theta+\delta(\underline{x}))d\theta - \int W(\theta)\pi_{\underline{x}}(\theta+\delta(\underline{x}))d\theta \qquad (1.3) \\ &= \tfrac{1}{2}\int [W(\theta+t') + W(\theta-t') - 2W(\theta)]\pi_{x}(\theta+ \delta(\underline{x}))d\theta, \end{aligned}$$

where $t' = t-\delta(\underline{x})$. Thus for convex functions W, the quantities in (1.3) are nonnegative, which implies (1.1).

In this paper we describe under mild regularity restrictions, the densities $p(x,\theta)$ and generalized priors $\lambda(\theta)$ for which the relation (1.2) holds. For these families a universal estimator exists for any CS set of loss functions.

Section 2 contains the major part of the proof of the main theorem, Theorem 2.1. In section 3 we discuss the case of a multivariate location parameter and the statistical properties of the distributions obtained. The solution of a functional equation of D'Alembert's type is needed to complete the proof of Theorem 2.1, and is given in section 4.

2. THE MAIN RESULT.

THEOREM 2.1 Let $\{p(x,\theta), x \in X, \theta \in \Theta\}$ be a family of probability densities given on a differentiable manifold X of dimension ρ where Θ is an open symmetric connected subset of $\mathbb{R}^m$. Assume that, for each $\theta \in \Theta$, $p(x,\theta)$ is a positive differentiable function of x. Assume also that $\rho \geq m$, and that $p(x,\theta)$ is continuous in θ for each fixed x, and that $\lambda(\theta)$ is a positive continuous function of θ. Suppose further that for some $n \geq 3$ and all $\underline{x} \in X^n$ (except for x from some nowhere dense set N) the function $\pi_{\underline{x}}(\theta) = \prod_1^n p(x_j,\theta)\lambda(\theta)$ is symmetric with respect to the point $\delta(\underline{x})$, where δ is a continuous function on $X^n \setminus N$ and $\delta(X^n \setminus N) = \Theta$. Then there exist non-negative integers $q_1,\ldots,q_R$ satisfying $q_1 + \ldots + q_R = m$ such that

$$\log p(x,\theta) + \frac{1}{n}\lambda(\theta) = \sum_{r=1}^{R} \sum_{i\leq i,k\leq q_r} d_{ik}^{(r)}(x)$$

$$\times [e^{\frac{1}{2}<\alpha_r,t-b(x)>} P_i^{(r)}(\frac{t-b(x)}{2}) - e^{-\frac{1}{2}<\alpha_r,t-b(x)>} P_i^{(r)}(\frac{b(x)-t}{2})]$$

$$\times [e^{\frac{1}{2}<\alpha_r,t-b(x)>} P_k^{(r)}(\frac{t-b(x)}{2}) - e^{-\frac{1}{2}<\alpha_r,t-b(x)>} P_k^{(r)}(\frac{t-b(x)}{2})].$$

Here the $P_k^{(r)}(-t) = -P_k^{(r)}(t)$, are polynomials with complex coefficients of degree $p_k^{(r)}$, and $p_k^{(1)} \leq 2q_1-1$, $k=1,\ldots,q_1$; $p_k^{(r)} \leq q_r-1$, $j=1,\ldots,q_r$, $r=2,\ldots,R$; $\alpha_1=0$, $\alpha_r \in \mathbb{C}$, $\alpha_r=0$, $r=2,\ldots,R$; and $d_{ik}^{(r)}(x) = d_{ki}^{(r)}(x)$, $1 \leq i,k \leq q_r$, $r=1,\ldots,R$, and $b(x)$ are real functions of x. The coefficients ${}_rF^{(k)}_{j_1\ldots j_m}$ of the polynomial $P_k^{(r)}$ satisfy the relations (2.7) and (2.8).

Proof. For $x \in X$ and $s \in \Theta$ define

$$\psi(x,s) = \log p(x,s) + \frac{1}{n}\log \lambda(s).$$

Then from the assumption that $\underline{x} \notin N$, it follows that

$$\sum_1^n \psi(x_j,\delta(\underline{x})-s) = \sum_1^n \psi(x_j,\delta(\underline{x})+s).$$

Let

$$w_\sigma(t) = (2\pi)^{-m/2} \prod_1^m \sigma_i^{-1} \exp\{-\tfrac{1}{2}\sum_1^m \frac{t_i^2}{\sigma_i^2}\},$$

and

$$\psi_\sigma(x,s) = \int \psi(x,s-t)w_\sigma(t)dt.$$

Then $\psi_\sigma(x,s)$ is infinitely differentiable in s and $\psi_\sigma(x,s) \to \psi(x,s)$ as $\sigma \to 0$.

Observe that for $\underline{x} \notin N$

$$\begin{aligned}\sum_1^n \psi_\sigma(x_j,\delta(\underline{x})-s) &= \sum_1^n \int \psi(x_j,\delta(\underline{x})-s-t)w_\sigma(t)dt \\ &= \sum_1^n \int \psi(x_j,\delta(\underline{x})+s+t)w_\sigma(t)dt \\ &= \sum_1^n \int \psi(x_j,\delta(\underline{x})+s-t)w_\sigma(t)dt \\ &= \sum_1^n \psi_\sigma(x_j,\delta(\underline{x})+s).\end{aligned} \tag{2.1}$$

For fixed $\underline{x} \notin N$, define

$$\overline{\psi}_\sigma(s) = \sum_{j=1}^n \psi_\sigma(x_j,\delta(\underline{x})+s).$$

We first claim that there exist non-zero numbers $\tau_1,\ldots,\tau_m$ such that

$$D(\sigma,\tau) = \det\{ \frac{1}{2\tau_i} \frac{\partial}{\partial s_k} [\overline{\psi}_\sigma(s+\tau_i e_i)-\overline{\psi}_\sigma(s-\tau_i e_i)]|_{s=0} \} = 0,$$

where $e_1,\ldots,e_m$ is a basis of $\mathbb{R}^m$ and $\tau = (\tau_1,\ldots,\tau_m)$. Indeed $D(\sigma,\tau)$ is an analytic function of σ for $\sigma \neq 0$, and the limit of $D(\sigma,\tau)$ as $\tau \to 0$ is $\det[\frac{\partial^2}{\partial s_i \partial s_k} \overline{\psi}_\sigma(s)|_{s=0}]$. This determinant does not vanish for small σ, since otherwise there exists an i, i=1,...,m, such that

$$\int [-1 + \frac{s_i^2}{\sigma_i^2}]\exp\{-\tfrac{1}{2} \sum_1^m \frac{s_i^2}{\sigma_i^2}\} \sum_1^n \psi(x_j,\delta(\underline{x})+s)ds = 0$$

for all σ. By known properties of the Laplace transform, this relation implies that $\frac{\partial}{\partial s_i} \sum_1^n \psi(x_j,\delta(\underline{x})+s) = 0$, which is impossible. Thus there exist linearly independent vectors $t_i = \tau_i e_i$, i=1,...,m, such that for all sufficiently small σ

$$\det\{\frac{\partial}{\partial s_k} \sum_1^n [\psi_\sigma(x_j,s+t_i) - \psi_\sigma(x_j,s-t_i)]\Big|_{s=\delta(\underline{x})}\} \neq 0. \qquad (2.2)$$

Denote $\phi(x,t) = (\psi_\sigma(x,t+t_1)-\psi_\sigma(x,t-t_1),\ldots,\psi_\sigma(x,t+t_m)-\psi_\sigma(x,t-t_m))'$.

(The symbol ' denotes transposition). It follows from (2.2) that, for $\underline{x} \notin N$, $\sum_1^n \psi(x_j,s)$ is a local isomorphism at the point $s=\delta(\underline{x})$. In other words, there exists a neighborhood V of $\delta(\underline{x})$ such that the restriction of $\sum_1^n \psi(x_j,s)$ to V establishes an isomorphism between V and an open subset of $\mathbb{R}^m$. Thus the relation

$$\sum_1^n \psi_\sigma(x_j,t+t_i) = \sum_1^n \psi_\sigma(x_j,t-t_i) \qquad (2.3)$$

for $i=1,\ldots,m$ and $t \in V$ implies that $t = \delta(\underline{x})$.

Note also that if $x = (x^1,\ldots,x^\rho)$, $\rho \geq m$, where the x^i are coordinates of x, then the matrix $(\frac{\partial}{\partial x^i}\ \phi(x,t))$ has rank m for all t. (If this rank were less than m for some t, an application of the rank theorem (cf. Narasimhan (1968)) shows that there exist two diffeomorphisms g_1 and g_2 such that $g_1 \cdot \phi \cdot g_2$ has the form $(x^1,\ldots,x^r,0,\ldots,0)$ where $r < m$. Then $\sum_1^n \phi(x_j,t)$ could not be a local isomorphism.)

Let $\Delta_t = \{\underline{x}:\ \underline{x} \in X^n/N,\ \delta(\underline{x}) = t\}$, so that $X/N = \bigcup_t \Delta_t$. Also let X_t be the projection of Δ_t on X, $X_t = \{x: \exists x_2,\ldots,x_n, \delta(x,x_2,\ldots,x_n)\} = t\}$. Then $X \setminus \bigcup_t X_t$ is nowhere dense. We next show that the set $T_x = \{t: x \in X_t\}$ contains a nonempty open set if T_x is nonempty.

For every $s \in T_x$, there exists, because of (2.3), a neighborhood $\mathcal{V}_s$ of s such that $\mathcal{V}_s \cap T_x = \{t: t \in \mathcal{V}_s, \phi(x,t)+\phi(x_2 t)+\ldots+\phi(x_n,t)\} = 0$ for some $x_2,\ldots,x_n\}$. The implicit function theorem and the proven fact concerning the rank of the matrix $(\frac{\partial}{\partial x^i}\phi(x,t))$ imply that $\mathcal{V}_s \cap T_x$ contains a neighborhood of s. Since T_x is nonempty except for x in a nowhere dense set, it follows that T_x contains a nonempty open set except for x in a nowhere dense set.

For fixed t and $x \in X_t$, the relation $\phi(x,t) = \phi(x_1,t)$ implies that $\psi_\sigma(x,t+s)-\psi_\sigma(x,t+s) = \psi_\sigma(x_1,t+s)-\psi_\sigma(x_1,t-s)$ for all s. In fact, for some $x_2,\ldots,x_n$

$$\phi(x,t) = -\sum_{2}^{n} \phi(x_j,t) = \phi(x_1,t)$$

and $\delta(x_1,x_2,\ldots,x_n) = t = \delta(x,x_2,\ldots,x_n)$. Therefore

$$\psi_\sigma(x,t+s)-\psi_\sigma(x,t-s) = -\sum_{2}^{n} [\psi_\sigma(x_j,t+s)-\psi_\sigma(x_j,t-s)] = \psi_\sigma(x_1,t+s)-\psi_\sigma(x_1,t-s).$$

Thus for every s and t there exists a real function g such that

$$\psi_\sigma(x,t+s)-\psi_\sigma(x,t-s) = g(\phi(x,t)).$$

It can be easily proven that g is continuous, and that

$$g(z_1)+\ldots+g(z_n) = 0$$

if $z_1+\ldots+z_n = 0$. Since $n \geq 3$ it follows from this equation that

$$g(z_1+z_2) = g(z_1)+g(z_2).$$

Because of the continuity of g there exists a vector $h \in \mathbb{R}^m$ such that $g(z) = \langle h,z \rangle$(cf. Aczel (1966) p.302). Thus we have proved that

$$\psi_\sigma(x,t+s)-\psi_\sigma(x,t-s) = \sum_{i=1}^{m} h_i(t,s)\Phi_i(x,t). \tag{2.4}$$

Note that (2.4) is true for all s and $x \in X_t$, where the set $T_x = \{t\colon x \in X_t\}$ contains a neighborhood except for x from a nowhere desne set.

It follows from (2.1) that for those $x_j \in X_t$ satisfying $\delta(x_1,\ldots,x_n) = t_0$ $\sum_{1}^{n} \psi_\sigma(x_j,t) = \sum_{1}^{n} \psi_\sigma(x_j,2t_0-t)$.

Hence

$$\sum_{1}^{m} h_i(t,s) \sum_{1}^{n} \phi_i(x_j,t) = \sum_{1}^{n}[\psi_\sigma(x_j,t+s)-\psi_\sigma(x_j,t-s)]$$

$$= \sum_{1}^{n}[\psi_\sigma(x_j,2t_0-t-s)-\psi_\sigma(x_j,2t_0-t+s)]$$

$$= -\sum_1^m h_i(2t_0-t,s) \sum_1^n \phi_i(x_j,2t_0-t)$$

$$= \sum_1^m h_i(2t_0-t,s) \sum_1^n \phi_i(x_j,t).$$

The last relation follows from the equalities

$$\sum_1^n \phi_i(x_j,2t_0-t) = \sum_1^n [\psi_\sigma(x_j,2t_0-t+t_i)-\psi_\sigma(x_j,2t_0-t-t_i)]$$

$$= -\sum_1^n [\psi_\sigma(x_j,t-t_i)-\psi_\sigma(x_j,t+t_i)] = -\sum_1^n \phi_i(x_j,t).$$

Thus

$$\sum_1^m [h_i(t,s)-h_i(2t_0-t,s)] \sum_1^n \phi_i(x_j,t) = 0.$$

This relation implies that $h_i(t,s) = h_i(2t_0-t,s)$ for all $i=1,\ldots,m$. In fact since the rank of the matrix $(\frac{\partial}{\partial x^i} \phi(x,t))$ is equal to m, one can find $\underline{x}^{(1)},\ldots,\underline{x}^{(m)} \in X^n/N$ such that the vectors $\sum_1^n \phi(\underline{x}_j^{(k)},t)$ are linearly independent. Therefore $h_i(t,s) = h_i(s)$, $i=1,\ldots,m$, and the functional equation (2.4) can be rewritten in the following form:

$$\psi_\sigma(x,t+s)-\psi_\sigma(x,t-s) = \sum_1^m h_i(s)\phi_i(x,t).$$

It is clear that if $\phi(x,t) = 0$, then $\psi_\sigma(x,t+s) = \psi_\sigma(x,t-s)$ for all s.

Let $\varphi_\sigma(x,t) = \psi_\sigma(x,t+b(x))$, where $\phi(x,b(x)) = 0$. Then $\varphi(x,-t) = \varphi(x,t)$ and

$$\varphi(x,t+s)-\varphi(x,t-s) = \sum_{i=1}^m h_i(s)\phi_i(x,t). \tag{2.5}$$

Now it follows from the proof of Theorem 4.1 that there exist nonnegative integers $q_1,\ldots,q_R$ $(q_1+\ldots+q_R=m)$, such that

$$\varphi(x,t) = \sum_{r=1}^R \sum_{1\le i,k\le q_r} d_{ik}^{(r)}(x)$$
$$\times [e^{\frac{1}{2}\langle\alpha_r,t-b(x)\rangle} P_i^{(r)}(\tfrac{t-b(x)}{2})-e^{-\frac{1}{2}\langle\alpha_r,t-b(x)\rangle} P_i^{(r)}(\tfrac{b(x)-t}{2})] \tag{2.6}$$
$$\times [e^{\frac{1}{2}\langle\alpha_r,t-b(x)\rangle} P_k^{(r)}(\tfrac{t-b(x)}{2})-e^{-\frac{1}{2}\langle\alpha_r,t-b(x)\rangle} P_k^{(r)}(\tfrac{b(x)-t}{2})],$$

where $P_k^{(r)}(t) = \sum_r F_{i_1 \ldots i_m} \frac{t_1^{i_1} \ldots t_m^{i_m}}{i_1! \ldots i_m!}$, $P_k^{(r)}(-t) = -P_k^{(r)}(t)$ are polynomials with complex coefficients of degree $p_k^{(r)}$, $p_k^{(1)} \leq 2q_1-1$; $p_k^{(r)} \leq q_r-1$, $r \geq 2$, $\alpha_1=0$, $\alpha_r \neq 0$, $r \geq 2$, $\alpha_r \in \mathbb{C}^m$; $d_{ik}^{(r)}(x) = d_{ki}^{(r)}(x)$, $i,k=1,\ldots,q_r$, $r=1,\ldots,R$. Independence of the polynomials $P_k^{(r)}$ and vectors α_r of x can be seen by substituting (2.6) in (2.5). Also, there exist complex numbers $B_{k_1 \ldots k_m}^{(ki)}$, such that

$$\Sigma\, B_{k_1 \ldots k_m}^{(ki)}\, {}_1F_{i_1 \ldots i_m}^{(i)} = {}_1F_{i_1+k_1 \ldots i_m+k_m}^{(k)}, \tag{2.7}$$

and for all x

$$\sum_\ell d_{i\ell}^{(1)} B_{k_1 \ldots k_m}^{(\ell k)} = \sum_\ell d_{k\ell}^{(1)}(x) B_{k_1 \ldots k_m}^{(\ell i)} \quad .$$

qor $r \geq 2$,

$${}_rF_{i_1 \ldots i_m}^{(\ell)} = \langle {}_rN_1^k \ldots {}_rN_m^m f_r, e_\ell \rangle, \tag{2.8}$$

where ${}_rN_i$ is a lower triangular matrix with zero diagonal, f_r is a fixed vector, e_ℓ are linearly independent vectors, and the matrices ${}_rN_i$ satisfy the relation (4.20). Letting σ go to zero completes the proof.

3. DISCUSSION. THE LOCATION PARAMETER CASE. It follows from Theorem 2.1 that any family of densities with a universal estimator has the form

$$p(x,\theta) = C(\theta)\exp\{\varphi(x,\theta)\},$$

where $C(\theta) = [\lambda(\theta)]^{1/n}$ and $\varphi(x,t)$ is given by the formula (2.6). Thus $p(x,\theta)$ belongs to the exponential family. When the prior density $\tilde{\lambda}(\theta)$ is from the conjugate family, i.e, has the form

$$\tilde{\lambda}(\theta) = [C(\theta)]^{-n}\exp\{\varphi(x_0,\theta)\} \quad x_0 \in X,$$

the posterior density is symmetric.

The most interesting example of the densities (3.1) arises when θ is a location parameter. Then $\Theta = \mathbb{R}^m = X$, μ is Lebesque measure, and

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

so that $C(\theta)$ = constant.

Theorem 2.1 implies that φ is symmetric with respect to some point θ_0. By shifting the initial density we can assume that $\theta = 0$. Also if $\varphi(c-t) = \varphi(c+t)$ for all t and some c, then $\varphi(t-c) = \varphi(t+c)$ for all t, implying that $c = 0$.

From the proof of the Theorem 4.1 it is clear that

$$\varphi(x-t) = \langle D(x)h(\frac{t-b(x)}{2}), h(\frac{t-b(x)}{2})\rangle,$$

where the matrix D(x) has elements $d_{ik}^{(r)}(x)$, and the vector function h arises from the polynomials $P_k^{(r)}(x)$. Thus for all t

$$\varphi(x-b(x)-t) = \varphi(x-b(x)+t),$$

so that $b(x) = x$, and

$$\varphi(t) = \langle D(x)h(t/2), h(t/2)\rangle.$$

It follows that $D(x) = D$ and all functions $d_{ik}^{(r)}(x)$ are constant. Note that $p(x-\theta)$ can be written

$$p(x-\theta) = C \exp\{\sum_{\ell=1}^{L} T_\ell(x) Q_\ell(\theta)\},$$

with linearly independent functions $\{T_\ell(x)\}$ and $\{Q_\ell(\theta)\}$, where the $Q_\ell(\theta)$ are polynomials of the form $\theta_1^{\ell_1}\cdots\theta_m^{\ell_m}$, $\ell_1+\ldots+\ell_m \le \max_{r,j} p_j^{(r)}$, or have one of the two forms $e^{\langle\alpha_r,\theta\rangle}\theta_1^{\ell_1}\cdots\theta_m^{\ell_m}$, $\ell_1+\ldots+\ell_m \le \max_j p_j^{(r)}$ or $e^{-\langle\alpha_r,\theta\rangle}\theta_1^{\ell_1}\cdots\theta_m^{\ell_m}$, $\ell_1+\ldots+\ell_m < \max_j p_j^{(r)}$. Therefore

$$L \le \binom{2\max_{r,j} p_j^{(r)}+m}{m} + 2\sum_{r=2}^{R}\binom{2\max_j p_j^{(r)}+m}{m}.$$

Another formula for the density p is

$$p(x) = C \exp\{\langle\exp[-x_1L_1-\ldots-x_mL_m]a,b\rangle\}.$$

Here $L_1,\ldots,L_m$ are commuting matrices of order L, a,b fixed vectors. To see this, note that because of (3.2), the linear space $\mathcal{L}$ spanned by all functions $\log p(\cdot-\theta)$, $\theta \in \mathbb{R}^m$, is finite-dimensional with functions $T_1(x),\dots,T_k(x)$ constituting its basis. Define the operator $M(\theta)$, $\theta \in \mathbb{R}^m$, acting in $\mathcal{L}$ by $M(\theta)\psi(\cdot) = \psi(\cdot-\theta)$, $\psi \in \mathcal{L}$. Then $M(\theta)$ is a finite-dimensional operator, and $M(\theta_1+\theta_2) = M(\theta_1)M(\theta_2)$. It follows that

$$M(\theta) = \exp\{-\theta_1 L_1-\dots-\theta_m L_m\},$$

where the matrices $L_1,\dots,L_m$ are commuting and have the order L. Therefore

$$\log p(x) = \langle M(x)a,b\rangle$$

for some vectors a, $b \in \mathcal{L}$.

Because of (2.9) the linear subspace $\mathcal{L}_0$ of $\mathcal{L}$ generated by the vectors $[\exp\{t_1L_1+\dots+t_mL_m\}-\exp\{-t_1L_1-\dots-t_mL_m\}]a$, $t_1,\dots,t_m \in \mathbb{R}^1$, is of dimension m. In other words, there are exactly m linearly independent vectors among $L_1^{k_1}\dots L_m^{k_m}a$, $k_1+\dots+k_m$ odd. Note that the subspace $\mathcal{L}_0$, which is spanned by latter vectors, is invariant for all operators $L_1^2,\dots,L_m^2$.

The symmetry condition $\varphi(-x) = \varphi(x)$ implies that

$$\langle[M(t)-M(-t)]a,b\rangle = 0,$$

i.e. the vector b belongs to the orthogonal complement of the space $\mathcal{L}_0$. The universal estimator $\delta(\underline{x})$ satisfies the relation

$$\langle[M(t)-M(-t)]a, \sum_1^n M^*(x_j-\delta(\underline{x})b\rangle = 0.$$

4. THE FUNCTIONAL EQUATION OF D'ALEMBERT'S TYPE

The theorem proved in this section gives a solution of the following functional equation:

$$\varphi(s+t)-\varphi(s-t) = \sum_{i=1}^m h_i(s)k_i(t). \qquad (4.1)$$

Here φ is a real continuous function defined on an open connected subset $\mathcal{D}$ (containing zero) of Euclidean space $\mathbb{R}^k$ and (4.1) holds for some continuous functions h_i, k_i, $i=1,\ldots,m$, and for all s, t, $s+t$, $s-t \in \mathcal{D}$. Without loss of generality it can be assumed that $h_i(s)$ and $k_i(t)$ are linearly independent. (Otherwise we can replace in (4.1) the linearly dependent functions with their representations as linear combinations relative to a basis of linearly independent functions, the result of which is an equation of similar form).

Let $<\alpha,\beta> = \sum_{c=1}^{m} \alpha_i\bar{\beta}_i$ denote the inner product of any two vectors α, $\beta \in \mathbb{C}^m$. With this notation, the equation (4.1) can be rewritten as

$$\varphi(s+t)-\varphi(s-t) = <h(s),k(t)>,$$

where $h'(s) = (h_1(s),\ldots,h_m(s))$ and $k'(t) = (k_1(t),\ldots,k_m(t))$.

Theorem 4.1 Let φ be an even continuous function defined on a symmetric open connected subset of $\mathbb{R}^k$, $\theta \in \mathcal{D}$, which satisfies (4.1) for some linearly independent functions h_i and k_i, $i=1,\ldots,m$. Then there exist nonnegative integers $q_1,\ldots,q_R$, $q_1+\ldots+q_R=m$, such that φ admits the following representation:

$$\varphi(s) = \sum_{j=1}^{q_1} d_j^{(1)}[P_j^{(1)}(s/2)]^2 + \sum_{r=2}^{R}\sum_{j=1}^{q_r} d_j^{(r)}[e^{\frac{1}{2}<\alpha_r,s>}P_j^{(r)}(s/2)-e^{-\frac{1}{2}<\alpha_r,s>}P_j^{(r)}(-s/2)]^2,$$

where the $P_j^{(r)}$, $P_j^{(r)}(-t) = -P_j^{(r)}(t)$ are polynomials with complex coefficients of degree $p_j^{(r)}$, $p_j^{(1)} \leq 2q_1-1$, $j=1,\ldots,q_1$; $p_j^{(r)} \leq q_r-1$, $j=1,\ldots,q_r$, $r=2,\ldots,R$; $d_j^{(r)} \in \mathbb{R}^1$, $j=1,\ldots,q_r$, $r=1,\ldots,R$; $\alpha_r \in \mathbb{C}^m$, $\alpha_r \neq 0$, $r=2,\ldots,R$. The coefficeints $F_{i_1\ldots i_m}^{(j)}$ of the polynomial $P_j^{(1)}$ satisfy the relations (4.11) for some constants $B_{k_1\ldots k_m}^{(kj)}$ satisfying condition (4.18). The coefficients $\Phi_{i_1\ldots i_m}^{(j)}$ of the polynomial $P_j^{(r)}$, $j=1,\ldots,q_r$, $r=2,\ldots,R$, satisfy the relation (4.19) with matrices $N_j^{(r)}$ satisfying condition (4.20). Every function φ of such a form is a solution of (4.1) for some functions k_i, h_i, $i=1,\ldots,m$.

Proof. The symmetry assumption $\varphi(-t) = \varphi(t)$ implies that $k(-t) = -k(t)$ and that

$$\langle h(s),k(t)\rangle = \langle h(t),k(s)\rangle.$$

Note that elements $\bar{t}_i$, $i=1,\ldots,m$, can be found such that the vectors $h(\bar{t}_i)$ are linearly independent (cf. for instance Aczel (1966) p. 201). If the matrix A is defined by the relation $k(\bar{t}_j) = Ah(\bar{t}_j)$, then $k(s) = A'h(s)$ and

$$\langle Ah(s),h(t)\rangle = \langle Ah(t),h(s)\rangle.$$

Hence $A'=A$. Since the functions $h_i(t)$ are linearly independent, A is non-singular, and

$$\varphi(s+t)-\varphi(s-t) = \langle Ah(s),h(t)\rangle.$$

Now for all s, t, u

$$\varphi(s+t+u)-\varphi(s+t-u) = \langle Ah(s+t),h(u)\rangle$$

and

$$\varphi(s-t+u)-\varphi(s-t-u) = \langle Ah(s-t),h(u)\rangle.$$

It follows from these relations that

$$\begin{aligned}\langle A[h(s+t)+h(s-t)]h(u)\rangle &= \varphi(s+t+u)-\varphi(s-t-u) + \varphi(s-t+u)-\varphi(s+t-u)\\ &= \langle Ah(s),h(t+u)\rangle + \langle Ah(s),h(-t+u)\rangle.\end{aligned}$$

Define matrices A(t) by the formula

$$h(\bar{t}_j+t) + h(\bar{t}_j-t) = 2A(t)h(\bar{t}_j), \quad j=1,\ldots,m.$$

Then

$$A[h(t+u) + h(-t+u)] = 2A'(t)Ah(u)$$

and

$$\langle Ah(s),A(t)h(u)\rangle = \langle h(s),A'(t)Ah(u)\rangle.$$

Thus for all t

$$AA(t) = A'(t)A \tag{4.3}$$

and

$$h(t+u) + h(-t+u) = 2A(t)h(u). \tag{4.4}$$

Note that $A(-t) = A(t)$ for all t. It is also clear that

$$\begin{aligned} 2A(s)A(t)h(u) &= A(s)[h(t+u)+h(-t+u)] \\ &= \tfrac{1}{2}[h(s+t+u)+h(-s+t+u)+h(s-t+u)+h(-s-t+u)] \\ &= [A(s+t)+A(s-t)]h(u). \end{aligned}$$

Because of this relation the matrices $A(t)$ satisfy D'Alembert's functional equation

$$A(s+t) + A(s-t) = 2A(s)A(t). \tag{4.5}$$

An immediate consequence of (4.5) is that all matrices $A(t)$ are commutative. It is known (cf. Suprunenko and Tyshkevich (1968) p. 16) that the whole space $\mathbb{R}^m$ can be represented as a direct sum of invariant subspaces Q_r, with respect to all $A(t)$, for $r=1,\ldots,R$. The irreducible parts of the restrictions $A(t)|Q_r$ are equivalent, while for $r\neq w$ the irreducible parts of $A(t)|Q_r$ and $A(t)|Q_w$ are not equivalent. Because of Shur's lemma (cf. for instance [7] pp. 4,8) all irreducible parts of $A(t)|Q_r$, $r=1,\ldots,R$, are one-dimensional operators. Hence all matrices $A(t)$ can be simultaneously reduced to the form $A(t) = T^{-1}B(t)T$, with complex matrices T and $B(t)$, the latter being of the form

$$B(t) = \begin{pmatrix} B_1(t) & 0 & \cdots & 0 \\ 0 & B_2(t) & \cdots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & & B_R(t) \end{pmatrix}.$$

Here $B_r(t)$, $r=1,\dots,R$, is a lower triangular matrix of dimension $q_r = \dim Q_r$ given by

$$B_r(t) = \begin{pmatrix} b^{(r)}(t) & 0 & \dots & 0 \\ b^{(r)}_{21}(t) & b^{(r)}(t) & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ b^{(r)}_{q_r 1}(t) & b^{(r)}_{q_r 2}(t) & \dots & b^{(r)}(t) \end{pmatrix},$$

where $b^{(r)}(t) \neq b^{(w)}(t)$, $r \neq w$, and $q_1+\dots+q_R = m$.

If $\tilde{A} = (T^{-1})^*AT^{-1}$, then $\tilde{A}^* = \tilde{A}$ and

$$\tilde{A}B(t) = B^*(t)\tilde{A}. \tag{4.6}$$

Therefore $\tilde{A}$ has the form

$$\tilde{A} = \begin{pmatrix} \tilde{A}_1 & 0 & \dots & 0 \\ 0 & \tilde{A}_2 & \dots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & & \tilde{A}_R \end{pmatrix},$$

with symmetric $q_r \times q_r$ matrices $\tilde{A}_r$, $r=1,\dots,R$. It follows from (4.4) that if $f(s) = Th(s)$, then

$$2B(t)f(s) = f(s+t)+f(s-t) \tag{4.7}$$

and $\langle Ah(s),h(t)\rangle = \langle \tilde{A}f(s),f(t)\rangle$. Clearly $f(-s) = -f(s)$.

Let $f'(s) = (f_1(s),\dots,f_R(s))$, where the vector-function f_r has dimension q_r, $r=1,\dots,R$. The relations (4.5) and (4.7) imply that

$$B_r(s+t) + B_r(s-t) = 2B_r(s)B_r(t) \tag{4.8}$$

and

$$2B_r(t)f_r(s) = f_r(s+t) + f_r(s-t), \quad r=1,\ldots,R. \tag{4.9}$$

It follows from (4.8) that

$$b^{(r)}(s+t) + b^{(r)}(s-t) = 2b^{(r)}(s)b^{(r)}(t).$$

All solutions of this D'Alembert's functional equation are known (cf. Kannappan (1968)) to be of the form

$$b^{(r)}(s) = \cosh(\alpha_1^{(r)}s_1+\ldots\alpha_m^{(r)}s_m) = \cosh\langle\alpha_r,s\rangle, \tag{4.10}$$

for some complex numbers $\alpha_1^{(r)},\ldots,\alpha_m^{(r)}$.

Consider first the case of vanishing α, say, $\alpha_1=0$, i.e. $b^{(1)}(s)=1$. Then $B_1(t) = I+N(t)$, where $N^q(t)=0$, $q=q_1$, and I is the identity matrix. It follows from (4.8) and (4.9) that

$$N(s+t)-2N(s)+N(s-t) = 2N(t)+2N(t)N(s)$$

and

$$f_1(s+t)-2f_1(s)+f_1(s-t) = 2N(t)f_1(s). \tag{4.11}$$

If one defines the operator $L(t)$ by the relation $L(t)f(\cdot)=f(\cdot-t)$, $t \in \mathbb{R}^m$, then (4.11) can be rewritten as

$$[L(t/2)-L(-t/2)]^2f_1(s) = 2N(t)f_1(s).$$

It follows by induction that

$$[L(t/2)-L(-t/2)]^{2q}f_1(s) = [2N(t)]^qf_1(s) = 0,$$

so that (cf. for example Aczel (1966) p. 130) each coordinate function $f_j^{(1)}(s)$ of $f_1(s)$ is a polynomial. It is easy to see that $f_j^{(1)}(s)$ is a polynomial of degree less than $2q_1$ (Actually, $f_j^{(1)}(s)$ is a polynomial of degree less than $2j$, $j=1,\ldots,q$.) Thus

$$f_j^{(1)}(s) = \sum_{i_\ell\leq 2q-1} F^{(j)}_{i_1\ldots i_m}\frac{s_1^{i_1}\ldots s_m^{i_m}}{i_1!\ldots i_m!},$$

where $s'=(s_1,\ldots,s_m)$. Since $f_1(-s)=-f_1(s)$, we have $F^{(j)}_{i_1\ldots i_m}=0$ if $\sum_1^q i_1$ is an even number. Analogously,

$$b^{(1)}_{kj}(s) = \sum_{k_\ell\leq 2q-2} B^{(kj)}_{k_1\ldots k_m} \frac{s_1^{k_1}\ldots s_m^{k_m}}{k_1!\ldots k_m!}, \qquad k > j,$$

and $B^{(kj)}_{k_1\ldots k_m}=0$ if $\sum_1^q k_\ell$ is an odd number or $\sum_1^q k_\ell = 0$.

Substituting these expressions for $f_1(s)$ and $N(t)$ into (4.10) shows that

$$\sum_j B^{(kj)}_{k_1\ldots k_m} F^{(j)}_{i_1\ldots i_m} = F^{(k)}_{i_1+k_1\ldots i_m+k_m}. \tag{4.12}$$

If $\tilde{A}_1 = \{\alpha_{ik}, 1\leq i,k\leq q\}$, $\alpha_{ik} = \bar{\alpha}_{ki}$, then (4.6) implies

$$\sum_\ell \alpha_{i\ell} B^{(\ell k)}_{k_1\ldots k_m} = \sum_\ell \alpha_{k\ell} B^{(\ell i)}_{k_1\ldots k_m}. \tag{4.13}$$

Note from (4.12) and (4.13) that if $k_j \geq \ell_j$, $j=1,\ldots,m$, then

$$\sum_{i,k} \alpha_{ik} F^{(i)}_{i_1\ldots i_m} F^{(k)}_{k_1\ldots k_m} = \sum_{i,k} \alpha_{ik} F^{(i)}_{i_1+\ell_1\ldots i_m+\ell_m} F^{(k)}_{k_1-\ell_1\ldots k_m-\ell_m}$$

Hence if $\sum_1^m (i_j+k_j) > 2m$, it follows that

$$\sum_{i,k} \alpha_{ik} F^{(i)}_{i_1\ldots i_m} F^{(k)}_{k_1\ldots k_m} = 0. \tag{4.14}$$

Indeed we can take $\sum_1^m i_j \geq 2m+1$, which will imply that $F^{(i)}_{i_1\ldots i_m} = 0$.

Now let us return to (4.10) and consider the case of non-zero α, say, $\alpha_2=\alpha\neq 0$, $q_2=p$, $B_2(t)=B(t)$. In this case, $B^2(t)-I$ is nonsingular for all t such that $\langle\alpha,t\rangle\neq 2\Pi ia$ for integer a. Therefore, there exists a t_0 and a nonsingular complex lower triangular matrix G such that $G^2=B^2(t_0)-I = \frac{1}{2}[B(2t_0)-I]$. Indeed

$$B^2(t)-I = 2[\sinh\langle\alpha,\tfrac{t}{2}\rangle]^2[I+M(t)],$$

where $M^p(t) = 0$. Thus we can take

$$G = \sqrt{2}\ \sin h\langle\alpha,\frac{t_0}{2}\rangle[I+\tfrac{1}{2}M(t_0) + \sum_{k=2}^{p-1} \frac{(-1)^{k+1}(2k-3)!!}{2^k k!} M^k(t_0)].$$

Clearly $G^2 = B^2(t_0)-I$, and G commutes with all matrices B(t).

Now let

$$G(t) = G^{-1}[B(t)(G-B(t_0)) + B(t+t_0)]$$
$$= B(t)-G^{-1}[B(t)B(t_0)-B(t+t_0)].$$

It is easy to check (cf. Kannapan (1968)) that

$$G(s+t) = G(s)G(t),$$

so that $G(s) = \exp\{s_1G_1+\ldots+s_mG_m\}$. Here the G_i are complex triangular matrices with diagonal elemetns all equal to $\alpha_i^{(2)}$, $i=1,\ldots,m$.

It follows from the definition of G(t) that

$$G(t)+G(-t) = 2B(t)-G^{-1}[2B(t)B(t_0)-B(t+t_0)-B(-t+t_0)] = 2B(t).$$

Since all matrices B(t) commute, the matrices G_i, $i=1,\ldots,m$, commute as well. From (4.6), one concludes that

$$\tilde{A}_2G_i^2 = [G_i^*]^2\tilde{A}_2.$$

The equation (4.7) implies that

$$f_2(s+t)+f_2(s-t) = 2\cos h(\sum_{i=1}^m t_iG_i)f_2(s).$$

Also

$$f_2(s+t)+f_2(t-s) = 2\cos h(\sum_{i=1}^m s_iG_i)f_2(t),$$

so that

$$f_2(s+t) = \cos h(\sum_{i=1}^m t_iG_i)f_2(s)+\cos h(\sum_{i=1}^m s_iG_i)f_2(t).$$

On the other hand, if $\langle\alpha,s+t\rangle \neq i\frac{2a+1}{2}\pi$, a an integer, then

$$f_2(s)+f_2(t) = 2\cos h(\tfrac{1}{2}\sum_{i=1}^m (s_i-t_i)G_i)f_2(\tfrac{s+t}{2})$$
$$= \cos h(\tfrac{1}{2}\sum_{i=1}^m (s_i-t_i)G_i)[\cos h(\tfrac{1}{2}\sum_{i=1}^m (s_i+t_i)G_i)]^{-1}f_2(s+t).$$

Therefore

$$\cosh\left(\tfrac{1}{2}\sum_{i=1}^{m}(s_i-t_i)G_i\right)\left[\cosh\left(\sum_{i=1}^{m}t_iG_i\right)f_2(s)+\cosh\left(\sum_{i=1}^{m}s_iG_i\right)f_2(t)\right]$$

$$=\cosh\left(\tfrac{1}{2}\sum_{i=1}^{m}(s_i+t_i)G_i\right)[f_2(s)+f_2(t)],$$

or

$$\left[\cosh\left(\tfrac{1}{2}\sum_{i=1}^{m}(s_i+t_i)G_i\right)+\cosh\left(\tfrac{1}{2}\sum_{i=1}^{m}(s_i-3t_i)G_i\right)\right]f_2(s)$$

$$+\left[\cosh\left(\tfrac{1}{2}\sum_{i=1}^{m}(3s_i-t_i)G_i\right)+\cosh\left(\tfrac{1}{2}\sum_{i=1}^{m}(s_i+t_i)G_i\right)\right]f(t)$$

$$=2\cosh\left(\tfrac{1}{2}\sum_{i=1}^{m}(s_i+t_i)G_i\right)[f_2(s)+f_2(t)].$$

Combining these relations one obtains

$$\left[\cosh\left(\tfrac{1}{2}\sum_1^m(s_i+t_i)G_i\right)-\cosh\left(\tfrac{1}{2}\sum_1^m(s_i-3t_i)G_i\right)\right]f_2(s)$$

$$+\left[\cosh\left(\tfrac{1}{2}\sum_1^m(s_i+t_i)G_i\right)-\cosh\left(\tfrac{1}{2}\sum_1^m(3s_i-t_i)G_i\right)\right]f_2(t)=0$$

or

$$\sinh\left(\tfrac{1}{2}\sum_1^m(s_i-t_i)G_i\right)\sinh\left(\sum_1^m t_iG_i\right)f_2(s)=\sinh\left(\sum_1^m s_iG_i\right)\sinh\left(\tfrac{1}{2}\sum_1^m(s_i-t_i)G_i\right)f_2(t).$$

This implies that

$$f_2(t)=\sinh\left(\sum_1^m t_iG_i\right)f,$$

where f is a constant vector.

Since $G_i=\alpha_i I+N_i$, where $N_i=N_i^{(2)}$ and $N_i^P=0$, $i=1,\ldots,m$, $\exp\{\sum_1^m t_iG_i\}=$ $\exp\{\sum_{i=1}^m t_i\alpha_i\}\sum\frac{t_1^{k_1}\ldots t_m^{k_m}}{k_1!\ldots k_m!}N_1^{k_1}\ldots N_m^{k_m}$. Clearly $N_1^{k_1}\ldots N_m^{k_m}=0$ if $\sum_1^m k_i\geq p$. Therefore

$$f_j^{(2)}(t)=\tfrac{1}{2}[\exp\{\sum_{i=1}^m\alpha_it_i\}P_j^{(2)}(t)-\exp\{-\sum_{i=1}^m\alpha_it_i\}P_j^{(2)}(-t)],\ j=1,\ldots,p,$$

where $P_j^{(2)}(t)$ is a polynomial of degreee less than p,

$$P_j^{(2)}(t)=\sum_{\Sigma k_i\leq p-1}\frac{t_1^{k_1}\ldots t_m^{k_m}}{k_1!\ldots k_m!}\langle N_1^{k_1}\ldots N_m^{k_m}f,e_j\rangle=\sum_{\Sigma k_i\leq p-1}\frac{t_1^{k_1}\ldots t_m^{k_m}}{k_1!\ldots k_m!}\phi^{(j)}_{k_1\ldots k_m}\quad(4.15)$$

the e_j are basis vectors, and

$$2\alpha_i\tilde{A}_2N_i + \tilde{A}_2N_i^2 = 2\alpha_iN_i^*\tilde{A}_2 + [N_i^*]^2\tilde{A}_2, \quad i=1,\dots,m. \qquad (4.16)$$

We can, finally, give the formula for $\varphi(s)$. Because of (4.2),

$$\varphi(2s) = \langle Ah(s),h(s)\rangle = \langle \tilde{A}f(s),f(s)\rangle = \sum_{1\le i,j\le q_1} \alpha_{ij}^{(1)}P_i^{(1)}(s)P_j^{(1)}(s)$$

$$+ \sum_{r=2}^{R} \sum_{1\le i,j\le q_r} \alpha_{ij}^{(r)}[e^{\alpha_1^{(r)}s_1+\dots+\alpha_m^{(r)}s_m}P^{(i)}(s)-e^{-\alpha_1^{(r)}s_1-\dots-\alpha_m^{(r)}s_m}\times P^{(i)}(-s)]$$

$$\times [e^{\alpha_1^{(r)}s_1+\dots+\alpha_m^{(r)}s_m}P^{(j)}(s)-e^{-\alpha_1^{(r)}s_1-\dots-\alpha_m^{(r)}s_m}P^{(j)}(-s)]. \qquad (4.17)$$

Here the coefficients $F^{(j)}_{i_1\dots i_m}$ of the polynomial $P_j^{(r)}$ satisfy (4.12) for some quantities $B^{(kj)}_{k_1\dots k_m}$ which satisfy condition (4.13) (with $\alpha_{i\ell}$ replaced by $\alpha_{i\ell}^{(1)}$. The coefficients $\Phi^{(j)}_{k_1\dots k_m}$ are defined by (4.15) where the matrices N_i satisfy (4.16) (with general index r instead of 2). Note that because of (4.14) the degree of the first term in (4.17) does not exceed $2q_1$.

The formula for $\varphi(s)$, given in the theorem, follows from (4.17) since $\tilde{A}_r = U_rD_rU_r^*$ with unitary matrix U_r and diagonal matrix D_r, $r=1,\dots,R$. In this case (4.13) must be changed to

$$d_i^{(1)}B^{(ik)}_{k_1\dots k_m} = d_k^{(1)}B^{(ki)}_{k_1\dots k_m}, \qquad (4.18)$$

and (4.15), (4.16) to

$$P_j^{(r)}(t) = \sum_{\Sigma k_i\le q_r-1} \frac{t_1^{k_1}\dots t_m^{k_m}}{k_1!\dots k_m!}\langle [N_1^{(r)}]^{k_1}\dots[N_m^{(r)}]^{k_m}f_r,e_j^{(r)}\rangle$$

$$= \sum_{\Sigma k_i\le q_r-1} \frac{t_1^{k_1}\dots t_m^{k_m}}{k_1!\dots k_m!}\ {}_r\phi_{k_1\dots k_m}, \qquad (4.19)$$

where $[N_i^{(r)}]^{q_r} = 0$, $i=1,\dots,m$, $r=2,\dots,R$, and

$$2\alpha_i^{(r)}D_rN_i^{(r)} + D_r[N_i^{(r)}]^2 = 2\bar{\alpha}_i^{(r)}N_i^{(r)*}D_r + [N_i^{(r)*}]^2D_r, \qquad (4.20)$$

where D_r is the diagonal matrix with elements $\{d_j^{(r)},\ j=1,\dots,q_r\}$.

It remains only to prove that every function φ of the form (4.17) satisfies (4.1). From our argument it follows that the matrix B(t) is nonsingular on an everywhere dense set. For such t it follows from (4.7) that

$$2f(t) = B^{-1}(t)f(2t),$$

and that

$$f(s) + f(t) = 2B(\tfrac{s-t}{2})f(\tfrac{s+t}{2}) = B(\tfrac{s-t}{2})B^{-1}(\tfrac{s+t}{2})f(s+t)$$

if $B(\frac{s+t}{2})$ is nonsingular.

It is clear that if

$$\varphi(2s) = <\tilde{A}f(s),f(s)>,$$

where $\tilde{A}$ satisfies (4.6), then

$$\varphi(2s) - \varphi(2t) = <\tilde{A}f(s+t),f(s-t)>. \qquad (4.21)$$

Indeed for s and t such that $B(\frac{s+t}{2})$ and $B(\frac{s-t}{2})$ are nonsingular matrices

$$<\tilde{A}f(s+t),f(s-t)>$$

$$= <\tilde{A}B(\tfrac{s+t}{2})B^{-1}(\tfrac{s-t}{2})[f(s)+f(t)],B(\tfrac{s-t}{2})B^{-1}(\tfrac{s+t}{2})[f(s)-f(t)]>$$

$$= <B^{-1*}(\tfrac{s-t}{2})B^*(\tfrac{s+t}{2})A[f(s)+f(t)],B(\tfrac{s-t}{2})B^{-1}(\tfrac{s+t}{2})[f(s)-f(t)]>$$

$$= <A[f(s)+f(t)],f(s)-f(t)> = <Af(s),f(s)> - <Af(t),f(t)>.$$

Sinc f is continuous, the relation (4.21) is true for all s and t. This completes the proof.

Corollary. If m=1, then the given solutions of (4.1) reduce to the known ones (cf. [1] p. 175): $\varphi(s) = a \cosh <\alpha,s> + b$, or $\varphi(s) = <\beta,s>^2 + d$.

If m=2, all solutions of (4.1) have one of the following forms: $\varphi(s) = \alpha_1 s_1^2 + \alpha_2 s_1 s_2 + \alpha_3 s_1^2$, $\alpha_1 \neq 0$; $\varphi(s) = \beta_1 s_1^2 + \beta_2 s_1 s_2 + \beta_3 s_2^2$; or $\varphi(s) = \gamma_1[\sinh(\delta_1 s_1+\delta_1 s_2)]^2 + \gamma_2(\xi_1 s_1+\xi_2 s_2)\sinh(\delta_1 s_1+\delta_1 s_2) \times \cosh(\delta_1 s_1+\delta_2 s_2) + \gamma_3(\xi_1 s_1+\xi_2 s_2)^2[\cosh(\delta_1 s_1+\delta_2 s_2)]^2$, where $\gamma_3\delta_1\xi_1 = 0$ and $\gamma_3\delta_2\xi_2 = 0$. This shows that the solutions of (4.1) corresponding to a certain m are not necessarily linear combinations of solutions corresponding to smaller m.

REFERENCES

[1] Aczel, J. (1966). Lectures on Functional Equations and Their Applications. Academic Press, New York.

[2] Deutsch, R. (1965). Estimation Theory, Prentice-Hall, Englewood Cliffs, New Jersey.

[3] Kannappan, P. (1968). The Functional Equations $f(xy) + f(xy^{-1}) = 2f(x)f(y)$ for Groups, Prec. Amer. Math. Soc. 19, 69-74.

[4] Narasimhan, R. (1968). Analysis on Real and Complex Manifolds, Masson & Cie, Editeur-Paris North-Holland Publishing Co., Amsterdam.

[5] Rukhin, A. L. (1978a). Strongly Symmetrical Families and Statistical Analysis of their Parameters. J. Soviet Math. 9, 886-910.

[6] Rukhin, A. L. (1978b). Universal Bayes Estimators, Ann. Statist. 6, 1345-1351.

[7] Suprunenko, D. A. and Tyshkevich, R. I. (1968). Commutative Matrices, Academic Press, New York, London.