Andrej Pázman

Probability distribution of the multivariate nonlinear least squares estimates

Kybernetika, Vol. 20 (1984), No. 3, 209--230

Persistent URL: http://dml.cz/dmlcz/125581

PROBABILITY DISTRIBUTION OF THE MULTIVARIATE NONLINEAR LEAST SQUARES ESTIMATES

ANDREJ PÁZMAN

The nonlinear regression model $y_i = \eta_i(\theta_1, \ldots, \theta_m) + \varepsilon_i$ with $(\varepsilon_1, \ldots, \varepsilon_N) \sim N(0, \sum)$ and with $\eta_i(.)$ twice continuously differentiable is considered. Under the assumption that the maximal curvature of the mean-values manifold $\{\boldsymbol{\eta}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in U\} \subset \mathbb{R}^N$ is bounded, an approximative probability density for the least squares estimates of $(\theta_1, \ldots, \theta_m)$ is proposed. This density depends on the first form (= the information matrix) and on the second form of the mean-values manifold (Eq. (9)). The level of approximation depends on the probability that the sample goes beyond the nearest center of curvature of the mean-values manifold and it is expressed in the paper (Theorem 1).

1. INTRODUCTION AND MAIN RESULTS

As in [4], let us consider the gaussian nonlinear regression model

$$\mathbf{y} = \boldsymbol{\eta}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{y} := (y_1, \ldots, y_N)'$ is the vector of observed variables, $\boldsymbol{\theta} := (\theta_1, \ldots, \theta_m)'$ is the vector of unknown parameters and $\boldsymbol{\varepsilon} := (\varepsilon_1, \ldots, \varepsilon_N)'$ is the vector of random observations errors. It is supposed that $\boldsymbol{\theta} \in U \subset \mathbb{R}^m$, U open, and that $\boldsymbol{\varepsilon}$ is distributed normally, $N(0, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}$ known and nonsingular. The functions $\eta_1, \ldots, \eta_N$ are defined and have continuous second order derivatives $\partial^2 \eta_k / \partial\theta_i \, \partial\theta_j$ on U. Finally, it is supposed that the vectors $\partial\boldsymbol{\eta}/\partial\theta_1, \ldots, \partial\boldsymbol{\eta}/\partial\theta_m$ are linearly independent for every $\boldsymbol{\theta} \in U$.

Eq. (1) could be also written in the more common form

$$y_i = \eta_{x_i}(\boldsymbol{\theta}) + \varepsilon_{x_i}\,; \quad (i = 1, \ldots, N)$$

where $x_1, \ldots, x_N$ are the points of the design of the experiment. The dependence of $\mathsf{E}(y_i)$ on x_i is of no importance in this paper, therefore we prefere the simpler Eq. (1).

Denote by $\langle \mathbf{a}, \mathbf{b} \rangle$, $\|\mathbf{a}\|$ the inner product and the norm defined by

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}' \, \boldsymbol{\Sigma}^{-1} \mathbf{b}\,, \quad \|\mathbf{a}\|^2 = \langle \mathbf{a}, \mathbf{a} \rangle\,.$$

The probability density of $\mathbf{y}$ is given by

(2) $$f(\mathbf{y} \mid \boldsymbol{\eta}(\boldsymbol{\theta})) = \frac{1}{(2\pi)^{N/2} \det^{1/2}(\boldsymbol{\Sigma})} \exp\{-\tfrac{1}{2}\|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\|^2\}.$$

The *least squares* (= l. s.) *estimate* for $\boldsymbol{\theta}$ is defined by

(3) $$\hat{\boldsymbol{\theta}} := \hat{\boldsymbol{\theta}}(\mathbf{y}) := \operatorname{Arg} \min_{\boldsymbol{\theta} \in U} \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\|^2$$

(if it exists). Hence the l. s. estimate $\hat{\boldsymbol{\theta}}$ is one of the solutions of the equations

$$\frac{\partial}{\partial \theta_i} \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\|^2 = 0\,; \quad (i = 1, \ldots, m),$$

or equivalently of

(4) $$\left\langle \mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}), \frac{\partial \boldsymbol{\eta}}{\partial \theta_i} \right\rangle = 0\,; \quad (i = 1, \ldots, m).$$

Geometrically it means that the vector $\mathbf{y} - \boldsymbol{\eta}(\hat{\boldsymbol{\theta}})$ is orthogonal to the *mean values manifold* (the set of potentially possible mean values).

$$\mathscr{E} := \{\boldsymbol{\eta}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in U\},$$

at the point $\boldsymbol{\eta}(\hat{\boldsymbol{\theta}})$. It means also that the vector $\mathbf{y}$ is in the hyperplane $\varkappa(\hat{\boldsymbol{\theta}})$ where

(5) $$\varkappa(\boldsymbol{\theta}) := \{\mathbf{z} : \mathbf{z} \in \mathbb{R}^N, \langle \mathbf{z} - \boldsymbol{\eta}(\boldsymbol{\theta}), \partial \boldsymbol{\eta}/\partial \theta_i \rangle = 0;\ (i = 1, \ldots, m)\}.$$

Let us denote by r the minimal radius of curvature of the manifold $\mathscr{E}$. More exactly, we denote by $r(\boldsymbol{\eta})$ the minimal radius of curvature of a geodesics which contains the point $\boldsymbol{\eta} \in \mathscr{E}$ (see Appendix for the properties of geodesics), and we define

$$r := \inf\{r(\boldsymbol{\eta}) : \boldsymbol{\eta} \in \mathscr{E}\}$$

AS 1: *We shall suppose in this paper that* $r > 0$, *i.e we consider models with bounded curvatures.*

Let $\chi^2_N(p_0)$ be the $(1 - p_0)$ quantile of the χ^2 p.d. with N degrees of freedom. If $\chi^2_N(p_0) = r^2$, we say that $(1 - p_0)$ is *the level of regularity* of the model.

It means that

$$\mathsf{P}_{\boldsymbol{\eta}}\{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \|\mathbf{y} - \boldsymbol{\eta}\| < r\} = 1 - p_0,$$

where $\mathsf{P}_{\boldsymbol{\eta}}$ is the p.d. with the density $f(\mathbf{y} \mid \boldsymbol{\eta})$.

We shall say that the regression model is *with a distant boundary* if for any expected $\boldsymbol{\eta} = \mathsf{E}(\mathbf{y})$ and any $\mathbf{y} \in \mathbb{R}^N$ such that $\|\mathbf{y} - \boldsymbol{\eta}\| < r$ there is a solution of Eq. (3). It means that we suppose that there is a set of expected values of the true vector $\boldsymbol{\theta}$, $U_0 \subset U$, which is sufficiently distant from "the boundary" of U.

AS 2: *We suppose in this paper that the model is with a distant boundary.*

The assumption AS 2 avoid to consider the "edges" of the manifold $\mathscr{E}$. Such an assumption is usually adopted also in the linear regression model

$$\mathbf{y} = \mathbf{F}\boldsymbol{\theta} + \boldsymbol{\varepsilon}\,; \quad (\boldsymbol{\theta} \in U) \tag{6}$$

($\mathbf{F}$ = a given $N \times m$ matrix). Here it is commonly supposed that $U = \mathbb{R}^m$ although in reality the values of $\theta_1, \ldots, \theta_m$ are always bounded à priori.

The model is called overlapping if for some $\mathbf{y} \in \mathbb{R}^N$ there are two solutions $\boldsymbol{\theta}^{(1)} \neq \boldsymbol{\theta}^{(2)}$ of Eqs. (4) such that

$$\|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}^{(1)})\| \leqq r\,, \quad \|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}^{(2)})\| \leqq r\,.$$

AS 3: *We suppose in this paper that the considered model is not overlapping.*

For any $\boldsymbol{\theta} \in U$ let us denote by

$$\{\mathbf{M}(\boldsymbol{\theta})\}_{ij} := \mathsf{E}_{\boldsymbol{\eta}(\boldsymbol{\theta})}\left\{\frac{\partial \ln f(\mathbf{y} \mid \boldsymbol{\eta}(\boldsymbol{\theta}))}{\partial \hat{\theta}_i}\,\frac{\partial \ln f(\mathbf{y} \mid \boldsymbol{\eta}(\boldsymbol{\theta}))}{\partial \hat{\theta}_j}\right\} = \left\langle \frac{\partial \boldsymbol{\eta}}{\partial \theta_i}, \frac{\partial \boldsymbol{\eta}}{\partial \theta_j}\right\rangle; \quad (i, j = 1, \ldots, m) \tag{7}$$

the (local) Fisher information matrix.

By

$$\mathbf{P}^{\boldsymbol{\theta}} := \sum_{k,l} \frac{\partial \boldsymbol{\eta}}{\partial \theta_k}\,\{\mathbf{M}^{-1}(\boldsymbol{\theta})\}_{kl}\,\frac{\partial \boldsymbol{\eta}}{\partial \theta_l}\,\boldsymbol{\Sigma}^{-1} \tag{8}$$

we denote the matrix of projection onto the plane which is tangent to the manifold $\mathscr{E}$ at the point $\boldsymbol{\eta}(\boldsymbol{\theta})$. Let us denote by $q(\boldsymbol{\theta} \mid \boldsymbol{\eta})$ the function

$$q(\boldsymbol{\theta} \mid \boldsymbol{\eta}) := \frac{\det\left[\left(\{\mathbf{M}(\boldsymbol{\theta})\}_{ij} + \left\langle(\mathbf{I} - \mathbf{P}^{\boldsymbol{\theta}})(\eta(\boldsymbol{\theta}) - \boldsymbol{\eta}), \frac{\partial^2 \boldsymbol{\eta}(\boldsymbol{\theta})}{\partial \theta_i\, \partial \theta_j}\right\rangle\right)_{i,j=1}^{m}\right]}{(2\pi)^{m/2} \det^{1/2} \mathbf{M}(\boldsymbol{\theta})} \times \exp\left\{-\tfrac{1}{2}\|\mathbf{P}^{\boldsymbol{\theta}}(\boldsymbol{\eta}(\boldsymbol{\theta}) - \boldsymbol{\eta})\|^2\right\}. \tag{9}$$

Let $f_{\boldsymbol{\theta}^\wedge}(\boldsymbol{\theta} \mid \boldsymbol{\eta})$ be the probability density* of the l.s. estimate $\hat{\boldsymbol{\theta}}$. The main result of the paper is expressed in Theorem 1 by the inequality

$$\left|\int_B f_{\boldsymbol{\theta}^\wedge}(\boldsymbol{\theta} \mid \boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\theta} - \int_B q(\boldsymbol{\theta} \mid \boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\theta}\right| \leqq 2p_0 \tag{10}$$

which is valid for every Borel set B which is a subset of $\{\boldsymbol{\theta} : \boldsymbol{\theta} \in U,\ \exists_{\mathbf{z} \in \mathbb{R}^N} \|\mathbf{z} - \boldsymbol{\eta}\| < r,\ \boldsymbol{\eta}(\boldsymbol{\theta}) = \boldsymbol{\eta}[\hat{\boldsymbol{\theta}}(\mathbf{z})]\}$ (= "the region of accessibility").

If follows that $q(\boldsymbol{\theta} \mid \boldsymbol{\eta})$ is an adequate approximative probability density of the l.s. estimates, the level of approximation being given by the level of regularity $(1 - p_0)$.

Especially, if $r \mapsto \infty$ (i.e. $p_0 \mapsto 0$), then we obtain the linear regression model (6). In that case

$$\partial^2 \boldsymbol{\eta}(\boldsymbol{\theta})/\partial \theta_i\, \partial \theta_j = 0\,, \quad \mathbf{P}^{\boldsymbol{\theta}^\wedge}[\boldsymbol{\eta}(\hat{\boldsymbol{\theta}}) - \boldsymbol{\eta}(\boldsymbol{\theta})] = \boldsymbol{\eta}(\hat{\boldsymbol{\theta}}) - \eta(\boldsymbol{\theta})\,,$$

$$\partial \eta_i(\boldsymbol{\theta})/\partial \theta_j = \{\mathbf{F}\}_{ij}\,, \quad \mathbf{M}(\boldsymbol{\theta}) = \mathbf{F}'\,\boldsymbol{\Sigma}^{-1}\,\mathbf{F}\,.$$

* From typographical reasons we use $\boldsymbol{\theta}^\wedge$, $\boldsymbol{\theta}^-$ instead of $\hat{\boldsymbol{\theta}}$, $\bar{\boldsymbol{\theta}}$ in superscript and subscript.

Thus from (9) we obtain the well known density

$$q(\hat{\boldsymbol{\theta}} \mid \boldsymbol{\eta}(\boldsymbol{\theta})) = \frac{\det^{1/2}(\mathbf{F}' \boldsymbol{\Sigma}^{-1} \mathbf{F})}{(2\pi)^{m/2}} \exp\{-\tfrac{1}{2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})' (\mathbf{F}' \boldsymbol{\Sigma}^{-1} \mathbf{F}) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\}.$$

2. CASE $m = 1, N = 2$ (HEURISTIC APPROACH)

To clarify the ideas we shall construct heuristically the probability density of the l.s. estimate in the special case $m = 1, N = 2$. Without restrictions on generality we shall suppose in this section that the parameter θ is the "natural parameter" (= the distance measured from some fixed point along the curve $\theta \in U \mapsto \boldsymbol{\eta}(\theta) \in \mathbb{R}^2$), i.e. that $\|d\boldsymbol{\eta}(\theta)/d\theta\| = 1$.

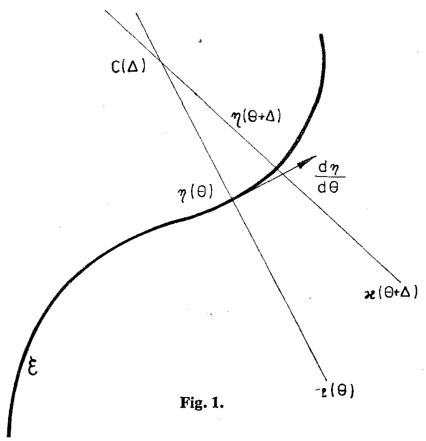

Fig. 1.

Denote by

(11)
$$\varphi(t) = [2\pi]^{-1/2} \exp\{-t^2/2\},$$
$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\, dt,$$

the (standarized) normal probability density function and the distribution function. By $S(\theta)$ we shall denote the half plane

$$S(\theta) := \left\{\mathbf{z} : \mathbf{z} \in \mathbb{R}^2, \left\langle \mathbf{z} - \boldsymbol{\eta}(\theta), \frac{d\boldsymbol{\eta}(\theta)}{d\theta} \right\rangle < 0 \right\}.$$

Take $\Delta > 0$. It can be seen from Fig. 1 that for a sufficiently small Δ the set

$$[S(\theta + \Delta) - S(\theta)] \cup [S(\theta) - S(\theta + \Delta)]$$

is the set of all points $\mathbf{y} \in \mathbb{R}^2$ which have a solution of Eq. (4) in the interval $(\theta, \theta + \Delta)$. Moreover, it can be seen from Fig. 1 that

$$(12) \qquad \mathsf{P}_{\boldsymbol{\eta}}[S(\theta + \Delta) - S(\theta)] - \mathsf{P}_{\boldsymbol{\eta}}[S(\theta) - S(\theta + \Delta)] = \\ = \mathsf{P}_{\boldsymbol{\eta}}[S(\theta + \Delta)] - \mathsf{P}_{\boldsymbol{\eta}}[S(\theta)]$$

where $\mathsf{P}_{\boldsymbol{\eta}}$ is the probability distribution of the sample $\mathbf{y}$ if $\boldsymbol{\eta}$ is its mean. Further evidently

$$\mathsf{P}_{\boldsymbol{\eta}}[S(\theta)] = \Phi\left[\left\langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, \frac{\mathrm{d}\boldsymbol{\eta}(\theta)}{\mathrm{d}\theta} \right\rangle\right].$$

We shall show that with $\Delta \to 0$ for $\mathbf{y} \in S(\theta + \Delta) - S(\theta)$ (resp. for $\mathbf{y} \in S(\theta) - S(\theta + \Delta)$) the solution of Eq. (4) is a relative minimum (resp. a relative maximum) of the function $\theta \in U \mapsto \|\mathbf{y} - \boldsymbol{\eta}(\theta)\|^2$.

To this purpose let us consider the second order derivative

$$(13) \qquad \frac{1}{2} \frac{\mathrm{d}^2}{\mathrm{d}\theta^2} \|\boldsymbol{\eta}(\theta) - \mathbf{y}\|^2 = 1 + \left\langle \boldsymbol{\eta}(\theta) - \mathbf{y}, \frac{\mathrm{d}^2\boldsymbol{\eta}}{\mathrm{d}\theta^2} \right\rangle.$$

The expression

$$\varrho_{\boldsymbol{\eta}}(\theta) := \left\|\frac{\mathrm{d}^2\boldsymbol{\eta}}{\mathrm{d}\theta^2}\right\|^{-1}$$

is the radius of curvature of the curve $\theta \in U \mapsto \boldsymbol{\eta}(\theta)$, and the point

$$\boldsymbol{\eta}(\theta) + \frac{\mathrm{d}^2\boldsymbol{\eta}}{\mathrm{d}\theta^2}$$

is its centre of curvature, as known from elementary differential geometry [1]. Let us denote by

$$\mathbf{e}_{\boldsymbol{\eta}}(\theta) := \varrho_{\boldsymbol{\eta}}(\theta) \frac{\mathrm{d}^2\boldsymbol{\eta}}{\mathrm{d}\theta^2}$$

the unit vector pointing from $\boldsymbol{\eta}(\theta)$ to the centre of curvature. This allows to rewrite Eq. (13) as

$$(14) \qquad \frac{1}{2} \frac{\mathrm{d}^2}{\mathrm{d}\theta^2} \|\boldsymbol{\eta}(\theta) - \mathbf{y}\|^2 = \varrho_{\boldsymbol{\eta}}^{-1}(\theta) \{\varrho_{\boldsymbol{\eta}}(\theta) + \langle \boldsymbol{\eta}(\theta) - \mathbf{y}, e_{\boldsymbol{\eta}}(\theta) \rangle\}.$$

As seen from Fig. 1, the point $C(\Delta)$ (= the point of intersection of $\varkappa(\theta)$ with $\varkappa(\theta + \Delta)$) tends to the centre of curvature if Δ tends to zero. For $\Delta \to 0$ form $\mathbf{y} \in S(\theta + \Delta) - S(\theta)$ it follows that $\langle \mathbf{y} - \boldsymbol{\eta}(\theta), \mathbf{e}_{\boldsymbol{\eta}}(\theta) \rangle < \varrho_{\boldsymbol{\eta}}(\theta)$, hence, according to Eq. (14), $\mathrm{d}^2/\mathrm{d}\theta^2 \|\boldsymbol{\eta}(\theta) - \mathbf{y}\|^2 > 0$. Moreover, as supposed, θ is the solution of Eq. (4), hence θ is a relative minimum of $\|\boldsymbol{\eta}(\theta) - \mathbf{y}\|^2$.

We proceed similarly in the case that $\mathbf{y} \in S(\theta) - S(\theta + \Delta)$. It follows that the limit

$$q(\theta \mid \boldsymbol{\eta}) := \lim_{\Delta \to 0} \frac{\mathsf{P}_{\boldsymbol{\eta}}[S(\theta + \Delta) - S(\theta)] - \mathsf{P}_{\boldsymbol{\eta}}[S(\theta) - S(\theta + \Delta)]}{\Delta}$$

is the probability density of the relative minima minus the probability density of the relative maxima of the function $\theta \in U \mapsto \|\boldsymbol{\eta}(\theta) - \mathbf{y}\|^2$.

From (12) it follows

$$q(\theta \mid \boldsymbol{\eta}) = \lim_{\Delta \to 0} \frac{\Phi\left[\left\langle \boldsymbol{\eta}(\theta + \Delta) - \boldsymbol{\eta}, \frac{\mathrm{d}\boldsymbol{\eta}(\theta + \Delta)}{\mathrm{d}\theta}\right\rangle\right] - \Phi\left[\left\langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, \frac{\mathrm{d}\boldsymbol{\eta}(\theta)}{\mathrm{d}\theta}\right\rangle\right]}{\Delta} =$$

$$(15) \qquad = \varphi\left(\left\langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, \frac{\mathrm{d}\boldsymbol{\eta}(\theta)}{\mathrm{d}\theta}\right\rangle\right) \frac{\mathrm{d}}{\mathrm{d}\theta}\left\langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, \frac{\mathrm{d}\boldsymbol{\eta}(\theta)}{\mathrm{d}\theta}\right\rangle .$$

If $\Delta \to 0$ then $[S(\theta) - S(\theta + \Delta)] \cap \{\mathbf{y} : \|\mathbf{y} - \boldsymbol{\eta}\| > r\} \to \emptyset$. Hence if we neglect the set of samples $\{\mathbf{y} : \mathbf{y} \in \mathbb{R}^2, \|\mathbf{y} - \boldsymbol{\eta}\| > r\}$, the probability of which is less than p_0, we can state that there are no relative maxima of $\|\boldsymbol{\eta}(\theta) - \mathbf{y}\|^2$ and that every relative minimum is an absolute minimum, i.e. it is the l.s. estimate $\hat{\theta}(\mathbf{y})$. Hence the expression in Eq. (15) is an approximative expression for the probability density of the l.s. estimate $\hat{\theta}$. To compare it with the expression in Eq. (9) we have just to use that in the special considered case $\mathbf{M}(\theta) = \|\mathrm{d}\boldsymbol{\eta}/\mathrm{d}\theta\|^2 = 1$ and that $\langle \mathrm{d}\boldsymbol{\eta}/\mathrm{d}\theta, \mathrm{d}^2\boldsymbol{\eta}/\mathrm{d}\theta^2\rangle = 0$.

The derivative

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left\langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, \frac{\mathrm{d}\boldsymbol{\eta}(\theta)}{\mathrm{d}\theta}\right\rangle_{\theta(\mathbf{y})} = \varrho^{-1}(\theta)\,[\varrho(\theta) + \langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, e_{\boldsymbol{\eta}}(\theta)\rangle]_{\theta(\mathbf{y})}$$

is positive (within our regularity assumptions). Hence, using the notation

$$(16) \qquad v(\theta) := \left\langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, \frac{\mathrm{d}\boldsymbol{\eta}(\theta)}{\mathrm{d}\theta}\right\rangle,$$

the approximative probability density $q(\theta \mid \boldsymbol{\eta})$ in (15) can be expressed as

$$q(\theta \mid \boldsymbol{\eta}) = (2\pi)^{-1/2} \exp\{-\tfrac{1}{2}v^2(\theta)\} \left|\frac{\mathrm{d}v(\theta)}{\mathrm{d}\theta}\right| .$$

It follows that the random variable $v(\hat{\theta})$ is (approximately) distributed $N(0, 1)$. Therefore, the interval

$$\left\{\theta : \left[\left\langle \boldsymbol{\eta}(\hat{\theta}) - \boldsymbol{\eta}(\theta), \frac{\mathrm{d}\boldsymbol{\eta}(\hat{\theta})}{\mathrm{d}\theta}\right\rangle\right]^2 < \chi_1^2(\beta)\right\}$$

is a confidence interval for the true value of θ, with the confidence level depending on β and on p_0.

Finally, let us compare the obtained probability density with the result in [4].

If $|\langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, \mathbf{e}_{\eta}(\theta)\rangle|$ is much smaller than $\varrho_{\eta}(\theta)$, then $|\mathrm{d}v(\theta)/\mathrm{d}\theta| \doteq 1$, and from (15) we obtain

$$q(\theta \mid \boldsymbol{\eta}) \doteq (2\pi)^{-1/2} \exp\{-\tfrac{1}{2}\langle \boldsymbol{\eta}(\theta) - \boldsymbol{\eta}, \mathrm{d}\boldsymbol{\eta}/\mathrm{d}\theta\rangle^2\}$$

which is the expression in Eq. (26) in [4] for the considered case that $\|\mathrm{d}\boldsymbol{\eta}/\mathrm{d}\theta\| = 1$.

3. THE MULTIVARIATE PROBABILITY DENSITY OF $\hat{\boldsymbol{\theta}}$

In this section we proceed to the general case of arbitrary m, N, $N > m$. We define by $\boldsymbol{\eta}$ the (fixed) mean of the sample $\mathbf{y}$. We denote by

$$(17) \qquad \mathscr{A}_{\eta} := \{\boldsymbol{\eta}[\hat{\boldsymbol{\theta}}(\mathbf{y})] : \|\mathbf{y} - \boldsymbol{\eta}\| < r\}$$

the region of accessibility (cf. Eq. (A 13) and the assumption AS 2 in Section 1). We are interested in the probability density $f_{\boldsymbol{\theta}^{\wedge}}(\boldsymbol{\theta} \mid \boldsymbol{\eta})$ of the l.s. estimate $\hat{\boldsymbol{\theta}}$. We shall show that it can be well approximated on the set $\{\boldsymbol{\theta} : \boldsymbol{\eta}(\boldsymbol{\theta}) \in \mathscr{A}_{\eta}\}$ by the function $q(\boldsymbol{\theta} \mid \boldsymbol{\eta})$ expressed in Eq. (9). The main aim of this section is to prove the following.

Theorem 1. Let B be a measurable subset of the set $\{\boldsymbol{\theta} : \boldsymbol{\theta} \in U, \boldsymbol{\eta}(\boldsymbol{\theta}) \in \mathscr{A}_{\eta}\}$. Then

$$\left|\int_B f_{\boldsymbol{\theta}^{\wedge}}(\boldsymbol{\theta} \mid \boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\theta} - \int_B q(\boldsymbol{\theta} \mid \boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\theta}\right| \leqq 2p_0 .$$

To prove Theorem 1 it is necessary to do a stepwise approximation of $f_{\boldsymbol{\theta}^{\wedge}}(\boldsymbol{\theta} \mid \boldsymbol{\eta})$ by $q(\boldsymbol{\theta} \mid \boldsymbol{\eta})$.

Take a fixed point $\bar{\boldsymbol{\theta}} \in U$ such that $\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) \in \mathscr{A}_{\eta}$. According to Proposition A 5, there is a neighbourhood of $\bar{\boldsymbol{\theta}}$, $U_{\bar{\boldsymbol{\theta}}} \subset U$, such that $\boldsymbol{\eta}[U_{\bar{\boldsymbol{\theta}}}] \subset \mathscr{A}_{\eta}$. As explained in Appendix, if a neighbourhood $V_{\bar{\boldsymbol{\theta}}} \subset U_{\bar{\boldsymbol{\theta}}}$ is adequately chosen, we can introduce new local coordinates $t_1, \ldots, t_m$ in $V_{\bar{\boldsymbol{\theta}}}$ and two sets of local coordinates $x_1, \ldots, x_N$ and $z_1, \ldots, z_N$ in the set $\mathscr{G}_{\bar{\boldsymbol{\theta}}} := \{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \mathbf{y} \in \varkappa(\boldsymbol{\theta}), \boldsymbol{\theta} \in V_{\bar{\boldsymbol{\theta}}}\}$ as follows. We take m geodesics in $\mathscr{E}$, $\gamma^{(1)}, \ldots, \gamma^{(m)}$ such that $\gamma^{(i)}(0) = \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})$; $(i = 1, \ldots, m)$ and that $\langle \dot{\gamma}^{(i)}(0), \dot{\gamma}^{(j)}(0)\rangle = 0$ if $i \neq j$ (cf. Appendix for geodesics in $\mathscr{E}$). The coordinates $t_1 := \tau_1(\boldsymbol{\theta}), \ldots, t_m := \tau_m(\boldsymbol{\theta})$ are defined by

$$(18) \qquad \langle \boldsymbol{\eta}(\boldsymbol{\theta}) - \gamma^{(i)}(t_i), \dot{\gamma}^{(i)}(t_i)\rangle = 0\,; \quad (i = 1, \ldots, m)$$

i.e. by projecting $\eta(\boldsymbol{\theta})$ onto the curves $\gamma^{(1)}, \ldots, \gamma^{(m)}$. We define further

$$(19) \qquad x_i = \xi_i(\mathbf{y}) := \tau_i[\boldsymbol{\theta}^*(\mathbf{y})]\,; \quad (i = 1, \ldots, m)\,,$$

where $\boldsymbol{\theta}^*(\mathbf{y})$ is the (unique) solution of Eqs. (4) which is in $V_{\bar{\boldsymbol{\theta}}}$. The coordinates $x_{m+1} = \xi_{m+1}(\mathbf{y}), \ldots, x_N = \xi_N(\mathbf{y})$ are complementary orthogonal coordinates defined by Eq. (A 22).

Projecting $\mathbf{y}$ onto $\gamma^{(1)}, \ldots, \gamma^{(m)}$, i.e. by the equations

$$(20) \qquad \langle \mathbf{y} - \gamma^{(i)}(z_i), \dot{\gamma}^{(i)}(z_i)\rangle = 0\,; \quad (i = 1, \ldots, m)$$

we define the coordinates $z_i = \zeta_1(\mathbf{y}), \ldots, z_m = \zeta_m(\mathbf{y})$. The coordinates $z_{m+1} = \zeta_{m+1}(\mathbf{y}), \ldots, z_N = \zeta_N(\mathbf{y})$ are again complementary orthogonal coordinates (cf. Eqs. (A 24)).

If $\mathbf{\Sigma} = \mathbf{I}$, and if $\boldsymbol{\theta}^*(\mathbf{y}) = \bar{\boldsymbol{\theta}}$, the coordinates $x_1, \ldots, x_N$ and $z_1, \ldots, z_N$ are essentially the same. In that case we have namely: $x_i = z_i$; $(i = 1, \ldots, N)$ and $\partial z_i/\partial x_j = 0$; $(i \neq j)$, $\partial z_i/\partial x_i = 1$ (cf. Eqs. (A 25) and Proposition A 7).

Let us introduce the notations

$$v_i(z_i) := \langle \gamma^{(i)}(z_i) - \boldsymbol{\eta}, \dot{\gamma}^{(i)}(z_i) \rangle \; ; \quad (i = 1, \ldots, m) \tag{21}$$

and

$$Q_i^{\boldsymbol{\theta}^-}(\mathbf{z}) := \{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, v_i[\zeta_i(\mathbf{y})] < v_i(\mathbf{z})\}$$

(cf. Eq. (A 17)).

Denote by $\zeta(\mathbf{y})$ the random vector

$$\zeta(\mathbf{y}) := (\zeta_1(\mathbf{y}), \ldots, \zeta_m(\mathbf{y}))\,,$$

and by $F_{\zeta}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m)$ its distribution function induced from the density of $\mathbf{y}$ given by Eq. (2), and restricted to the set $\mathscr{G}_{\boldsymbol{\theta}^-}$. Because the functions $v_1, \ldots, v_m$ defined in Eqs. (21) are increasing (see Proposition A 3), the increase of $F_{\zeta}^{\boldsymbol{\theta}^-}$ is given by

$$\Delta_{\varepsilon_1}^{(1)} \ldots \Delta_{\varepsilon_m}^{(m)} F_{\zeta}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m) = \mathsf{P}_{\boldsymbol{\eta}}\Big[\bigcap_{i=1}^{m} (Q_i^{\boldsymbol{\theta}^-}(z_i) - Q_i^{\boldsymbol{\theta}^-}(z_i - \varepsilon_i))\Big]\,. \tag{22}$$

Here we used the notation

$$\Delta_{\varepsilon_k}^{(k)} h(z_1, \ldots, z_m) := h(z_1, \ldots, z_m) - h(z_1, \ldots, z_{k-1}, z_k - \varepsilon_k, z_{k+1}, \ldots, z_m)\,,$$

(cf. [6], chpt. IV. 3). The density of $\zeta(\mathbf{y})$ is then

$$f_{\zeta}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m) := \frac{\partial^m}{\partial z_1 \ldots \partial z_m} F_{\zeta}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m) = \tag{23}$$

$$= \lim_{\varepsilon_1 \to 0} \ldots \lim_{\varepsilon_m \to 0} \frac{\mathsf{P}_{\boldsymbol{\eta}}\Big[\bigcap_{i=1}^{m} (Q_i^{\boldsymbol{\theta}^-}(z_i) - Q_i^{\boldsymbol{\theta}^-}(z_i - \varepsilon_i))\Big]}{\varepsilon_1 \ldots \varepsilon_m}\,.$$

Denote by $g(z_{m+1}, \ldots, z_N \mid z_1, \ldots, z_m)$ the conditional probability density of $\zeta_{m+1}(\mathbf{y}), \ldots, \zeta_N(\mathbf{y})$ (induced again from $f(\mathbf{y} \mid \boldsymbol{\eta})$ in Eq. (2)). The joint density $f_{\zeta}^{\boldsymbol{\theta}^-} \,.\, (z_1, \ldots, z_m)\, g(z_{m+1}, \ldots, z_N \mid z_1, \ldots, z_m)$ is transformed by the mapping (the change of coordinates) $(z_1, \ldots, z_N) \mapsto (x_1, \ldots, x_N)$ into the joint density of $(\xi_1(\mathbf{y}), \ldots, \xi_N(\mathbf{y}))$. Denote by $f_{\xi}^{\boldsymbol{\theta}^-}(x_1, \ldots, x_m)$ the corresponding marginal density of the random vector

$$\xi(\mathbf{y}) := (\xi_1(\mathbf{y}), \ldots, \xi_m(\mathbf{y}))\,.$$

Then finally, according to Eqs. (19),

$$f_{\boldsymbol{\theta}^\wedge}(\bar{\boldsymbol{\theta}} \mid \boldsymbol{\eta}) = f_{\xi}^{\boldsymbol{\theta}^-}(x_1, \ldots, x_m)\, \left|\det\left(\{\partial \tau_i/\partial \theta_j\}_{i,j=1}^{m}\right)\right| \tag{24}$$

This complicated way to deduce $f_{\boldsymbol{\theta}^\wedge}(\bar{\boldsymbol{\theta}} \mid \boldsymbol{\eta})$ from $f(\mathbf{y} \mid \boldsymbol{\eta})$ was chosen to make easy the comparison with $q(\bar{\boldsymbol{\theta}} \mid \boldsymbol{\eta})$. Namely, we shall show that $q(\bar{\boldsymbol{\theta}} \mid \boldsymbol{\eta})$ can be deduced in an analogical way, but from the distribution

$$\tilde{F}_{\xi}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m) := \mathsf{P}_{\boldsymbol{\eta}}[\bigcap_{i=1}^{m} S_i^{\boldsymbol{\theta}^-}(z_i)] \tag{25}$$

where

$$S_i^{\boldsymbol{\theta}^-}(z_i) := \{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \langle \mathbf{y} - \gamma^{(i)}(z_i), \dot{\gamma}^{(i)}(z_i)\rangle < 0\} \tag{26}$$

(cf. Eq. (A 16)). The corresponding density is given by

$$\tilde{f}_{\xi}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m) = \lim_{\varepsilon_1 \to 0} \ldots \lim_{\varepsilon_m \to 0} \frac{\mathsf{P}_{\boldsymbol{\eta}}[\bigcap_{i=1}^{m} (S_i^{\boldsymbol{\theta}^-}(z_i) - S_i^{\boldsymbol{\theta}^-}(z_i - \varepsilon_i))]}{\varepsilon_1 \ldots \varepsilon_m} \tag{27}$$

Again, the joint density $\tilde{f}_{\xi}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m)\, g(z_{m+1}, \ldots, z_N \mid z_1, \ldots, z_m)$ is transformed by the coordinate mapping $(z_1, \ldots, z_N) \mapsto (x_1, \ldots, x_N)$ into a joint density of $(\xi_1(\mathbf{y}), \ldots$ $\ldots, \xi_N(\mathbf{y}))$. Denote by $\tilde{f}_{\xi}^{\boldsymbol{\theta}^-}(x_1, \ldots, x_m)$ the corresponding marginal distribution of $\xi(\mathbf{y})$. We have the following important auxiliary proposition

Proposition 1. Let be $\boldsymbol{\Sigma} = \boldsymbol{I}$.
Then

a) $$f_{\xi}^{\boldsymbol{\theta}^-}(0) = f_{\zeta}^{\boldsymbol{\theta}^-}(0), \tilde{f}_{\xi}^{\boldsymbol{\theta}^-}(0) = \tilde{f}_{\zeta}^{\boldsymbol{\theta}^-}(0)$$

b) $$\tilde{f}_{\xi}^{\boldsymbol{\theta}^-}(0) = \frac{1}{(2\pi)^{m/2}} \exp\{-\tfrac{1}{2}(\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \boldsymbol{\eta})' \sum_{i=1}^{m} \dot{\gamma}^{(i)}(0)\, \dot{\gamma}^{(i)\prime}(0)\, (\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \boldsymbol{\eta})\} \times \times \prod_{i=1}^{m} [1 + (\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \boldsymbol{\eta})' \ddot{\gamma}^{(i)}(0)] \tag{28}$$

c) $$q(\bar{\boldsymbol{\theta}} \mid \boldsymbol{\eta}) = \tilde{f}_{\zeta}^{\boldsymbol{\theta}^-}(0) \left|\det(\{\partial\tau_i(\bar{\boldsymbol{\theta}})/\partial\theta_j\}_{i,j=1}^{m})\right| , \tag{29}$$

$$f_{\boldsymbol{\theta}^\wedge}(\bar{\boldsymbol{\theta}} \mid \boldsymbol{\eta}) = f_{\xi}^{\boldsymbol{\theta}^-}(0) \left|\det(\{\partial\tau_i(\bar{\boldsymbol{\theta}})/\partial\theta_j\}_{i,j=1}^{m})\right| . \tag{30}$$

Proof. The statement a) follows from Eqs. (A 25) and from Proposition A 7 in Appendix.

From Eqs. (25), (26) and (21) it follows that

$$\tilde{F}_{\xi}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m) = \int_{-\infty}^{v_1(z_1)} \ldots \int_{-\infty}^{v_m(z_m)} \frac{1}{(2\pi)^{m/2} \det^{1/2} \mathbf{K}} \exp\{-\tfrac{1}{2}\boldsymbol{u}'\mathbf{K}^{-1}\boldsymbol{u}\}\, du_1 \ldots du_m ,$$

where

$$\{\mathbf{K}\}_{ij} := \langle \dot{\gamma}^{(i)}(z_i), \dot{\gamma}^{(j)}(z_j)\rangle .$$

Therefore the corresponding density is

$$\tilde{f}_{\xi}^{\boldsymbol{\theta}^-}(z_1, \ldots, z_m) = \frac{1}{(2\pi)^{m/2} \det^{1/2} \mathbf{K}} \exp\{-\tfrac{1}{2} \sum_{i,j=1}^{m} v_i(z_i)\, \{\mathbf{K}^{-1}\}_{ij}\, v_j(z_j)\} \prod_{i=1}^{m} (\mathrm{d}v_i(z_i)/\mathrm{d}z_i) . \tag{31}$$

From (31) we obtain the expression in Eq. (28).

The equality in Eq. (30) follows directly from Eq. (24). It remains to prove Eq. (29). First we can state that

$$\mathbf{P}^{\theta^-} = \sum_{l=1}^{m} \dot{\gamma}^{(l)}(0)\, \dot{\gamma}^{(l)\prime}(0)\,. \tag{32}$$

To verify (32), put $\mathbf{P}^{\theta^-}$ according to Eq. (8) (for $\mathbf{\Sigma} = \mathbf{I}$), and multiply Eq. (32) by $\partial\boldsymbol{\eta}'(\bar{\boldsymbol{\theta}})/\partial\theta_q$ from the left, and by $\partial\boldsymbol{\eta}(\bar{\boldsymbol{\theta}})/\partial\theta_h$ from the right ($q, h = 1, \ldots, m$).

In the right side of Eq. (29) let us express $\tilde{f}_{\xi}^{\theta^-}(0)$ using Eq. (28), and $\partial\tau_i(\bar{\boldsymbol{\theta}})/\partial\theta_j$ using Eq. (A 27). We can write, according to Eq. (32),

$$[\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \boldsymbol{\eta}]' \sum_{i=1}^{m} \dot{\gamma}^{(i)}(0)\, \dot{\gamma}^{(i)\prime}(0)\, [\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \boldsymbol{\eta}] = \|\mathbf{P}^{\theta^-}[\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \boldsymbol{\eta}]\|^2\,. \tag{33}$$

Using Eq. (32) again, we obtain

$$\begin{aligned}
&\prod_{i=1}^{m} [1 + (\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \boldsymbol{\eta})'\, \ddot{\gamma}^{(i)}(0)]\, |\det(\{\partial\tau_i(\bar{\boldsymbol{\theta}})/\partial\theta_j\}_{i,j=1}^{m})| = \\
&= \det\left(\left\{\frac{\partial}{\partial t_i} [(\gamma^{(i)}(t_i) - \boldsymbol{\eta})' \sum_{k=1}^{m} \dot{\gamma}^{(k)}(t_k)\, \dot{\gamma}^{(k)\prime}(t_k)]_{t=0}\, \frac{\partial\boldsymbol{\eta}(\bar{\boldsymbol{\theta}})}{\partial\theta_j}\right\}_{i,j=1}^{m}\right) = \\
&= \det\left(\left\{\mathbf{M}_{ij}(\bar{\boldsymbol{\theta}}) + [\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \boldsymbol{\eta}]' (\mathbf{I} - \mathbf{P}^{\theta^-})\, \frac{\partial^2 \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})}{\partial\theta_i\, \partial\theta_j}\right\}_{i,j=1}^{m}\right) \times \\
&\qquad \times |\det^{-1}(\{\partial\tau_i(\bar{\boldsymbol{\theta}})/\partial\theta_j\}_{i,j=1}^{m})|\,.
\end{aligned} \tag{34}$$

From Eq. (A 27) we have

$$\det^2(\{\partial\tau_i(\bar{\boldsymbol{\theta}})/\partial\theta_j\}_{i,j=1}^{m}) = \det \mathbf{M}(\bar{\boldsymbol{\theta}})\,. \tag{35}$$

The validity of Eq. (29) follows from Eqs. (33)–(35). □

From Proposition 1 it follows that the comparison of $f_{\hat{\theta}}(\bar{\boldsymbol{\theta}} \mid \boldsymbol{\eta})$ with $q(\bar{\boldsymbol{\theta}} \mid \boldsymbol{\eta})$, needed in Theorem 1, reduces to the comparison of $f_{\xi}^{\theta^-}(0)$ with $\tilde{f}_{\xi}^{\theta^-}(0)$.

Proposition 2. For sufficiently small $\varepsilon_1 > 0, \ldots, \varepsilon_m > 0$ we have the inequality

$$\begin{aligned}
&|\Delta_{\varepsilon_1}^{(1)} \ldots \Delta_{\varepsilon_k}^{(k)}[F_{\xi}^{\theta^-}(z_1, \ldots, z_m) - \tilde{F}_{\xi}^{\theta^-}(z_1, \ldots, z_m)]| \leq \\
&\leq \mathsf{P}_{\boldsymbol{\eta}}[(\mathbb{R}^N - W_r) \bigcap_{i=1}^{m} (Q_i^{\theta^-}(z_i) - Q_i^{\theta^-}(z_i - \varepsilon_i))] + \\
&+ \mathsf{P}_{\boldsymbol{\eta}}[(\mathbb{R}^N - W_r) \bigcap_{i=1}^{m} (S_i^{\theta^-}(z_i) - S_i^{\theta^-}(z_i - \varepsilon_i))
\end{aligned} \tag{36}$$

where

$$W_r := \{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \|\mathbf{y} - \boldsymbol{\eta}\| < r\}\,.$$

Proof. From (25) we obtain

$$\Delta_{\varepsilon_1}^{(1)} \dots \Delta_{\varepsilon_m}^{(m)} \tilde{F}_{\zeta}^{\theta^-}(z_1, \dots, z_m) = \mathsf{P}_{\eta}[(\mathbb{R}^N - W_r) \bigcap_{i=1}^{m} (S_i^{\theta^-}(z_i) - S_i^{\theta^-}(z_i - \varepsilon_i))] + $$
$$+ \mathsf{P}_{\eta}[W_r \bigcap_{i=1}^{m} (S_i^{\theta^-}(z_i) - S_i^{\theta^-}(z_i - \varepsilon_i))] \,.$$

Analogically, from (22) we obtain

$$\Delta_{\varepsilon_1}^{(1)} \dots \Delta_{\varepsilon_m}^{(m)} F_{\zeta}^{\theta^-}(z_1, \dots, z_m) = \mathsf{P}_{\eta}[(\mathbb{R}^N - W_r) \bigcap_{i=1}^{m} (Q_i^{\theta^-}(z_i) - Q_i^{\theta^-}(z_i - \varepsilon_i))] + $$
$$+ \mathsf{P}_{\eta}[W_r \bigcap_{i=1}^{m} (Q_i^{\theta^-}(z_i) - Q_i^{\theta^-}(z_i - \varepsilon_i))] \,.$$

Hence, we prove the inequality (36) if we use that, according to Proposition A 6,

$$W_r \bigcap_{i=1}^{m} [Q_i^{\theta^-}(z_i) - Q_i^{\theta^-}(z_i - \varepsilon_i)] = W_r \bigcap_{i=1}^{m} [S_i^{\theta^-}(z_i) - S_i^{\theta^-}(z_i - \varepsilon_i)] \,,$$

for $\varepsilon_1, \dots, \varepsilon_m$ sufficiently small. □

Proof of Theorem 1. We shall write $\boldsymbol{\theta}$ instead of $\bar{\boldsymbol{\theta}}$ in this proof. Without lack of generality we shall do the proof for $\boldsymbol{\Sigma} = \mathbf{I}$.

According to Proposition 1 we have

$$D := \left| \int_B f_{\boldsymbol{\theta}^\wedge}(\boldsymbol{\theta} \mid \boldsymbol{\eta}) \, \mathrm{d}\boldsymbol{\theta} - \int_B q(\boldsymbol{\theta} \mid \boldsymbol{\eta}) \, \mathrm{d}\boldsymbol{\theta} \right| \leqq$$
$$\leqq \int_B |f_{\zeta}^{\boldsymbol{\theta}}(0) - \tilde{f}_{\zeta}^{\boldsymbol{\theta}}(0)| \, |\det(\{\partial \tau_i / \partial \theta_j\}_{i,j=1})| \, \mathrm{d}\boldsymbol{\theta} \,.$$

Hence, using Proposition 2, we obtain

$$(37) \qquad D \leqq \int_B \frac{\partial^m}{\partial z_1 \dots \partial z_m} \{\mathsf{P}_{\eta}[(\mathbb{R}^N - W_r) \bigcap_{i=1}^{m} S_i^{\boldsymbol{\theta}}(z_i)] + $$
$$+ \mathsf{P}_{\eta}[(\mathbb{R}^N - W_r) \bigcap_{i=1}^{m} Q_i^{\boldsymbol{\theta}}(z_i)]\}_{\mathbf{z}=\mathbf{0}} \, |\det(\{\partial \tau_i / \partial \theta_j\}_{i,j=1}^m)| \, \mathrm{d}\boldsymbol{\theta} =$$
$$= \int_B \frac{\partial^m}{\partial z_1 \dots \partial z_m} \int_{\mathbb{R}^N - W_r} \prod_{i=1}^{m} [\chi(\mathbf{y}; S_i^{\boldsymbol{\theta}}(z_i)) + \chi(\mathbf{y}; Q_i^{\boldsymbol{\theta}}(z_i))] \times$$
$$\mathrm{d}\mathsf{P}_{\eta}(\mathbf{y})|_{\mathbf{z}=\mathbf{0}} \, |\det(\{\partial \tau_i / \partial \theta_j\}_{i,j=1}^m)| \, \mathrm{d}\boldsymbol{\theta}$$

where $\chi(\mathbf{y}; T)$ denotes the indicator of a set T.

For fixed $\mathbf{y}$, $\boldsymbol{\theta}$ the functions $z_i \mapsto \chi(\mathbf{y}; S_i^{\boldsymbol{\theta}}(z_i))$; $(i = 1, \dots, m)$ have unit jumps at $\mathbf{z} = \mathbf{0}$ iff $\boldsymbol{\theta} = \boldsymbol{\theta}^*(\mathbf{y})$. As a consequence, from (37) it follows

$$D \leqq 2 \int_{\mathbb{R}^N - W_r} \chi(\boldsymbol{\theta}^*(\mathbf{y}); B) \, \mathrm{d}P_{\eta}(\mathbf{y}) \leqq 2p_0 \,.$$ □

Corollary. Let A be an arbitrary measurable subset of U. Then

$$\left|\int_A f_{\theta^\wedge}(\theta \mid \eta)\,\mathrm{d}\theta - \int_{A\cap\mathscr{B}_\eta} q(\theta \mid \eta)\,\mathrm{d}\theta\right| \leqq 3p_0$$

where $\mathscr{B}_\eta := \{\theta : \eta(\theta) \in \mathscr{A}_\eta\}$.

Proof. We have

$$\left|\int_A f_{\theta^\wedge}(\theta \mid \eta)\,\mathrm{d}\theta - \int_{A\cap\mathscr{B}_\eta} f_{\theta^\wedge}(\theta \mid \eta)\,\mathrm{d}\theta\right| \leqq \int_{\mathbb{R}^N - W_r} f(\mathbf{y} \mid \eta)\,\mathrm{d}\mathbf{y} = p_0 . \qquad \square$$

4. CONFIDENCE REGIONS FOR $\hat{\theta}$

Let us choose for every $\bar{\theta} \in U_0$ a set $I_{\bar{\theta}} \subset \{\theta : \eta(\theta) \in \mathscr{A}_{\eta(\bar{\theta})}\}$ such that

$$\int_{I_{\bar{\theta}}} q(\theta \mid \eta(\bar{\theta}))\,\mathrm{d}\theta \geqq 1 - \beta .$$

Then, according to Theorem 1,

$$\int_{I_{\bar{\theta}}} f_{\theta^\wedge}(\theta \mid \eta(\bar{\theta}))\,\mathrm{d}\theta \geqq 1 - \beta - 2p_0 .$$

Hence the set

$$\mathscr{I}_{\theta^\wedge} := \{\theta : \theta \in U_0, \hat{\theta} \in I_\theta\}$$

is a confidence region with the level of significance equal at least to $1 - \beta - 2p_0$.

Theorem 2. If for every $\theta \in U$ there is a (differentiable) orthonormal basis $\mathbf{l}_1(\theta), \ldots$ $\ldots, \mathbf{l}_m(\theta)$ of the tangent space to $\mathscr{E}$ at the point $\eta(\theta)$, such that

$$\frac{\partial \mathbf{l}_i'(\theta)}{\partial \theta_j}\,\mathbf{l}_k(\theta) = 0 ; \quad (i, j, k = 1, \ldots, m) \tag{38}$$

then we can take

$$\mathscr{I}_{\theta^\wedge} = \{\theta : \theta \in U_0, \|\mathbf{P}^\theta[\eta(\hat{\theta}) - \eta(\theta)]\|^2 < \chi_m^2(\beta)\}$$

where $\chi_m^2(\beta)$ is the $(1 - \beta)$ quantile of the χ^2 probability distribution with m degrees of freedom.

Proof. Take $\mathbf{\Sigma} = \mathbf{I}$. The expression in Eq. (9) can be written as

$$q(\theta \mid \eta) = \frac{1}{(2\pi)^{m/2}} \exp\left(-\tfrac{1}{2}\|\mathbf{P}^\theta[\eta(\theta) - \eta]\|^2\right) \frac{\det\left(\left\{\frac{\partial \eta'}{\partial \theta_j} \frac{\partial}{\partial \theta_k}(\mathbf{P}^\theta[\eta(\theta) - \eta])\right\}_{j,k=1}^m\right)}{\det^{1/2}\left(\left\{\frac{\partial \eta'}{\partial \theta_j} \frac{\partial \eta}{\partial \theta_k}\right\}_{j,k=1}^m\right)} . \tag{39}$$

Define

$$v_i(\boldsymbol{\theta}) := \boldsymbol{l}_i'(\boldsymbol{\theta})\,[\boldsymbol{\eta}(\boldsymbol{\theta}) - \boldsymbol{\eta}]\,; \quad (i = 1, \ldots, m)\,.$$

Evidently

$$\sum_{i=1}^{m} v_i^2(\boldsymbol{\theta}) = \|\mathbf{P}^{\theta}[\boldsymbol{\eta}(\boldsymbol{\theta}) - \boldsymbol{\eta}]\|^2\,. \tag{40}$$

Further, from (38) it follows that

$$\frac{\partial v_i(\boldsymbol{\theta})}{\partial \theta_k} = \boldsymbol{l}_i'(\boldsymbol{\theta})\,\frac{\partial}{\partial \theta_k}\,(\mathbf{P}^{\theta}[\boldsymbol{\eta}(\boldsymbol{\theta}) - \boldsymbol{\eta}])\,.$$

Hence

$$\frac{\partial \boldsymbol{\eta}'}{\partial \theta_j}\,\frac{\partial}{\partial \theta_k}\,(\mathbf{P}^{\theta}[\boldsymbol{\eta}(\boldsymbol{\theta}) - \boldsymbol{\eta}]) = \sum_{i=1}^{m} \frac{\partial \boldsymbol{\eta}'}{\partial \theta_j}\,\boldsymbol{l}_i\,\boldsymbol{l}_i'\,\frac{\partial}{\partial \theta_k}\,(\mathbf{P}^{\theta}[\boldsymbol{\eta}(\boldsymbol{\theta}) - \boldsymbol{\eta}]) = \sum_{i=1}^{m} \frac{\partial \boldsymbol{\eta}'}{\partial \theta_j}\,\boldsymbol{l}_i\,\frac{\partial v_i}{\partial \theta_k}\,. \tag{41}$$

From Eqs. (39)–(41) we obtain that

$$q(\boldsymbol{\theta} \mid \boldsymbol{\eta}) = \frac{1}{(2\pi)^{m/2}}\,\exp\left(-\tfrac{1}{2}\sum_i v_i^2(\boldsymbol{\theta})\right)\left|\det\left(\left\{\frac{\partial v_i(\boldsymbol{\theta})}{\partial \theta_j}\right\}_{i,j=1}^{m}\right)\right|,$$

hence the random vector $(v_1, \ldots, v_m)$ is distributed $N(0, \mathbf{I})$. The needed statement follows. □

Example 1. Take $m = 1$,

$$\boldsymbol{l}(\theta) = \frac{\mathrm{d}\boldsymbol{\eta}(\theta)/\mathrm{d}\theta}{\|\mathrm{d}\boldsymbol{\eta}(\theta)/\mathrm{d}\theta\|}\,.$$

Then

$$0 = \frac{\mathrm{d}}{\mathrm{d}\theta}\,(\boldsymbol{l}'\boldsymbol{l}) = 2\,\frac{\mathrm{d}\boldsymbol{l}'}{\mathrm{d}\theta}\,\boldsymbol{l}\,,$$

hence the assumption (38) is valid.

Example 2. Take $\mathscr{E}$ a subset of the cylinder

$$\{\mathbf{z} : \mathbf{z} \in \mathbb{R}^N,\ z_1^2 = 1\}\,.$$

Evidently Eq. (38) can be satisfied.

APPENDIX A

In this section we present some necessary geometrical statements. We start by some definitions.

A (regular) *curve* in U is a mapping

$$\mathbf{g} : t \in (a, b) \mapsto \mathbf{g}(t) \in U$$

such that the vector of second order derivatives $\mathrm{d}^2\mathbf{g}/\mathrm{d}t^2$ exists and it is continuous on (a, b). To the curve $\mathbf{g}$ we can associate a curve

$$\gamma : t \in (a, b) \mapsto \gamma(t) \in \mathscr{E}$$

according to

$$\gamma(t) = \boldsymbol{\eta}[\mathbf{g}(t)] . \tag{A 1}$$

The curve γ is called a *geodesics in the manifold* $\mathscr{E}$ (and correspondingly $\mathbf{g}$ is called a *geodesics in* U) iff

a) the parameter t is normed so that

$$\left\| \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right\| = 1 ; \quad (t \in (a, b)) \tag{A 2}$$

b) the *vector of curvature* $\mathrm{d}^2\gamma/\mathrm{d}t^2$ is orthogonal to $\mathscr{E}$ i.e.

$$\left\langle \frac{\mathrm{d}^2\boldsymbol{\eta}[\mathbf{g}(t)]}{\mathrm{d}t^2}, \frac{\partial\boldsymbol{\eta}[\mathbf{g}(t)]}{\partial\theta_i} \right\rangle = 0 ; \quad (i = 1, \ldots, m) \tag{A 3}$$

As known from differential geometry [3, 7], every nonzero solution $\mathbf{g}$ of the differential equations (A 3) (= the Euler-Lagrange equations) is a geodesics in U. Moreover, for every point $\bar{\boldsymbol{\theta}} \in U$ and every nonzero vector $\mathbf{u} \in \mathbb{R}^m$ there is a geodesics $\mathbf{g}$ such that for some $\bar{t}$

$$\mathbf{g}(\bar{t}) = \bar{\boldsymbol{\theta}} . \quad \mathrm{d}\mathbf{g}(\bar{t})/\mathrm{d}t \approx \mathbf{u} . \tag{A 4}$$

Correspondingly, to every point $\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) \in \mathscr{E}$ and to every unit vector $\mathbf{w} \in \mathbb{R}^N$ which is tangent to $\mathscr{E}$ at $\boldsymbol{\eta}(\bar{\boldsymbol{\theta}})$ (i.e. $\mathbf{w}$ is a linear combination of the vectors $\partial\boldsymbol{\eta}(\bar{\boldsymbol{\theta}})/\partial\theta_1, \ldots$ $\ldots, \partial\boldsymbol{\eta}(\bar{\boldsymbol{\theta}})/\partial\theta_m$) there is a geodesics γ such that

$$\gamma(\bar{t}) = \boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) , \quad \mathrm{d}\gamma(\bar{t})/\mathrm{d}t = \mathbf{w} . \tag{A 5}$$

We shall use the abbreviated notations

$$\dot{\gamma}(\bar{t}) = \mathrm{d}\gamma(t)/\mathrm{d}t|_{t=\bar{t}}$$
$$\ddot{\gamma}(\bar{t}) = \mathrm{d}^2\gamma(t)/\mathrm{d}t^2|_{t=\bar{t}}$$

We denote further

$$\varrho_\gamma(t) := \|\ddot{\gamma}(t)\|^{-1} \tag{A 6}$$
$$\mathbf{e}_\gamma(t) := \ddot{\gamma}(t)\,\varrho_\gamma(t)$$

the *radius of curvature* and the unit vector oriented from the point $\gamma(t)$ toward the centre of curvature. By

$$\varkappa_\gamma(t) := \{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \langle \gamma(t) - \mathbf{y}, \dot{\gamma}(t) \rangle = 0\} \tag{A 7}$$

we denote the hyperplane orthogonal to the curve γ.

Denote by $B(\boldsymbol{\theta}, \mathbf{y})$ (or simply by B) the $m \times m$ symmetric matrix with the entries

$$B_{ij}(\boldsymbol{\theta}, \mathbf{y}) := \left\langle \frac{\partial\boldsymbol{\eta}}{\partial\theta_i}, \frac{\partial\boldsymbol{\eta}}{\partial\theta_j} \right\rangle + \left\langle \boldsymbol{\eta}(\boldsymbol{\theta}) - \mathbf{y}, \frac{\partial^2\boldsymbol{\eta}}{\partial\theta_i\,\partial\theta_j} \right\rangle . \tag{A 8}$$

Proposition A 1. Let be $\mathbf{y} \in \mathbb{R}^N$ and let $\bar{\boldsymbol{\theta}}$ be a solution of Eqs. (4). $B(\bar{\boldsymbol{\theta}}, \mathbf{y})$ is positive semidefinite (= p.s.) iff for every geodesics γ such that $\gamma(\bar{t}) = \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})$ for some $\bar{t}$, we have the inequality

$$\text{(A 9)} \qquad \langle \mathbf{y} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}}), \mathbf{e}_\gamma(\bar{t}) \leqq \varrho_\gamma(\bar{t}) \,.$$

There is the equality sign in (A 9) for some geodesics γ iff $\det B(\bar{\boldsymbol{\theta}}, \mathbf{y}) = 0$.

Proof. Let $\gamma = \boldsymbol{\eta} \circ \boldsymbol{g}$ be a geodesics and let us denote $\mathbf{c} := \dot{\boldsymbol{g}}(t)$. From (A 8) we obtain

$$\text{(A 10)} \qquad \mathbf{c}'\mathbf{B}\mathbf{c} = 1 - \langle \mathbf{y} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}}), \mathbf{e}_\gamma(\bar{t})\rangle\, \varrho_\gamma^{-1}(\bar{t}) \,.$$

i) If $\mathbf{B}$ is p.s. then from Eq. (A 10) follows Eq. (A 9). Conversely, if Eq. (A 9) is valid for every geodesics γ then Eq. (A 10) implies that $\mathbf{c}'\,\mathbf{B}\mathbf{c} \geqq 0$ for every $\mathbf{c} \in \mathbb{R}^n$, such that $\mathbf{c} = \dot{\boldsymbol{g}}(t)$ for some geodesics $\boldsymbol{g}$. That means, according to Eqs. (A 2) and (A 4), $\mathbf{c}'\,\mathbf{B}\mathbf{c} \geqq 0$ for every $\mathbf{c}$ which is a solution of

$$\text{(A 11)} \qquad \mathbf{c}'\,\mathbf{M}(\bar{\boldsymbol{\theta}})\,\mathbf{c} = 1 \,.$$

Since $\mathbf{M}(\bar{\boldsymbol{\theta}})$ is positive definite it follows that $\mathbf{B}$ is p.s.

ii) If $\det \mathbf{B} = 0$ then $\mathbf{B}\mathbf{c} = 0$ for some $\mathbf{c} \in \mathbb{R}^m$, $\mathbf{c} \neq \boldsymbol{0}$. Leg $\boldsymbol{g}$ be the geodesics in U, such that $\boldsymbol{g}(\bar{t}) = \bar{\boldsymbol{\theta}}$ and that $\dot{\boldsymbol{g}}(\bar{t}) \approx \mathbf{c}$. Take $\gamma = \boldsymbol{\eta} \circ \boldsymbol{g}$. From (A 10) it follows that

$$\langle \mathbf{y} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}}), \mathbf{e}_\gamma(\bar{t})\rangle = \varrho_\gamma(\bar{t}) \,.$$

Conversely, if this equality is valid for some geodesics $\gamma = \boldsymbol{\eta} \circ \boldsymbol{g}$ then, according to (A 10), $\mathbf{c}'\,\mathbf{B}\mathbf{c} = 0$ for some $\mathbf{c} \neq \boldsymbol{0}$. Since $\mathbf{B}$ is p.s. there is a matrix $\mathbf{A}$ such that $\mathbf{B} = \mathbf{A}'\mathbf{A}$. Therefore $\|\mathbf{A}\mathbf{c}\|^2 = \mathbf{c}'\,\mathbf{B}\mathbf{c} = 0$. Thus $\mathbf{B}\mathbf{c} = 0$, and $\det \mathbf{B} = 0$. □

Corollary A 1. Let $\bar{\boldsymbol{\theta}}$ be a solution of Eqs. (4) and let $\|\mathbf{y} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})\| < r$. Then $\bar{\boldsymbol{\theta}}$ is the l.s. estimate $\hat{\theta}(\mathbf{y})$.

Proof. We have

$$\langle \mathbf{y} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}}), \mathbf{e}_\gamma(\bar{t})\rangle \leqq \|\mathbf{y} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})\| < r \leqq \varrho_\gamma(\bar{t}) \,.$$

Hence the matrix with entries

$$\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\, \|\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) - \mathbf{y}\|^2 = 2\,\mathbf{B}_{ij}(\bar{\boldsymbol{\theta}}, \mathbf{y})$$

is p.d. and $\bar{\boldsymbol{\theta}}$ is a relative minimum. The equality $\bar{\boldsymbol{\theta}} = \hat{\theta}(\mathbf{y})$ then follows from the assumption AS 3. □

Let us fix a point $\boldsymbol{\eta} \in \mathscr{E}$. Let us denote

$$\text{(A 12)} \qquad W_r := \{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \|\mathbf{y} - \boldsymbol{\eta}\| < r\} \,,$$

$$\overline{W}_r : \{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \|\mathbf{y} - \boldsymbol{\eta}\| \leqq r\} \,.$$

Cf. Eqs. (5), (A 7) and (17) for the definitions of $\varkappa(\boldsymbol{\theta})$, $\varkappa_\gamma(t)$ and $\mathscr{A}_{\boldsymbol{\eta}}$. From Corollary

A 1 and from the assumptions AS 2, AS 3 it follows that we can write

(A 13) $$\mathscr{A}_\eta = \{\boldsymbol{\eta}(\boldsymbol{\theta}) : \boldsymbol{\theta} \in U, \underset{\mathbf{z} \in \mathbb{R}^N}{\exists} \mathbf{z} \in \varkappa(\boldsymbol{\theta}), \|\mathbf{z} - \boldsymbol{\eta}\| < r, \|\mathbf{z} - \boldsymbol{\eta}(\boldsymbol{\theta})\| < r\} .$$

Proposition A 2. If $\mathbf{y} \in W_r \cap \varkappa(\bar{\boldsymbol{\theta}})$ and $\boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) \in \mathscr{A}_\eta$ then $\|\mathbf{y} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})\| < r$.

Proof. According to (A 13) there is a point $\mathbf{z} \in \varkappa(\bar{\boldsymbol{\theta}}) \cap W_r$ such that $\|\mathbf{z} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})\| < r$. Suppose that $\|\mathbf{y} - \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})\| \geqq r$. Consider the N-dimensional open sphere

$$\mathscr{S} := \{\mathbf{w} : \mathbf{w} \in \mathbb{R}^N, \|\mathbf{w} - \mathbf{c}\| < r\}$$

which is tangent to $\mathscr{E}$ at the point $\boldsymbol{\eta}(\bar{\boldsymbol{\theta}})$ and is such that $\mathbf{c}$ is on the straight line connecting $\mathbf{y}$ with $\mathbf{z}$. Evidently $\mathbf{c}$ is on the abscissa with the endpoints $\mathbf{y}, \mathbf{z}$, hence $\|\mathbf{c} - \boldsymbol{\eta}\| < r$. It follows that $\boldsymbol{\eta} \in \mathscr{S}$. But $\mathscr{S} \cap \mathscr{A}_\eta = \emptyset$. This is a contradiction to $\boldsymbol{\eta} \in \mathscr{A}_\eta$. □

Proposition A 3. Let γ be a geodesics and $\gamma(t) \in \mathscr{A}_\eta$ for some t. Then

$$\frac{\mathrm{d}^2 \|\gamma(t) - \boldsymbol{\eta}\|^2}{\mathrm{d}t^2} > 0 .$$

Proof. Take $\boldsymbol{\theta} \in U$ such that $\boldsymbol{\eta}(\boldsymbol{\theta}) = \gamma(t)$. Denote by $\boldsymbol{\eta}_\theta$ the point of projection of $\boldsymbol{\eta}$ onto $\varkappa(\boldsymbol{\theta})$. From Eqs. (5) and (A 3) we obtain

(A 14) $$\langle \boldsymbol{\eta}_\theta - \boldsymbol{\eta}, \ddot{\gamma}(t) \rangle = 0$$

According to (A 13) take $\mathbf{y} \in \mathbb{R}^N$ such that $\|\mathbf{y} - \boldsymbol{\eta}\| < r$, $\|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\| < r$, $\mathbf{y} \in \varkappa(\boldsymbol{\theta})$. We have $\|\boldsymbol{\eta} - \boldsymbol{\eta}_\theta\| \leqq \|\boldsymbol{\eta} - \mathbf{y}\| < r$, hence from Proposition A 2 we obtain $\|\boldsymbol{\eta}_\theta - \gamma(t)\| < r$. Therefore using (A 14) we can write

$$\frac{1}{2} \frac{\mathrm{d}^2}{\mathrm{d}t^2} \|\gamma(t) - \boldsymbol{\eta}\|^2 = 1 + \langle \gamma(t) - \boldsymbol{\eta}_\theta, \ddot{\gamma}(t) \rangle \geqq 1 - \|\boldsymbol{\eta}_\theta - \gamma(t)\| / \varrho_\gamma(t) > 0 . \quad \square$$

Proposition A 4. Let be $\boldsymbol{\eta}(\boldsymbol{\theta}^{(1)}) \in \mathscr{A}_\eta$, $\boldsymbol{\eta}(\boldsymbol{\theta}^{(2)}) \in \mathscr{A}_\eta$. Then $\varkappa(\boldsymbol{\theta}^{(1)}) \cap \varkappa(\boldsymbol{\theta}^{(2)}) \cap W_r = \emptyset$.

Proof. According to the assumption AS 3, $\mathbf{z} \in \varkappa(\boldsymbol{\theta}^{(1)}) \cap \varkappa(\boldsymbol{\theta}^{(2)})$ implies $\|\mathbf{z} - \boldsymbol{\eta}(\boldsymbol{\theta}^{(1)})\| > r$ or $\|\mathbf{z} - \boldsymbol{\eta}(\boldsymbol{\theta}^{(2)})\| > r$. Hence, according to Proposition A 2, $\mathbf{z} \notin W_r$. □

Proposition A 2. Let γ be a geodesics, $\gamma(\bar{t}) = \boldsymbol{\eta}(\bar{\boldsymbol{\theta}}) \in \mathscr{A}_\eta$. Then there is a neighbourhood of $\bar{\boldsymbol{\theta}}$, $U_{\bar{\theta}} \subset U$, such that $\boldsymbol{\eta}(U_{\bar{\theta}}) \subset \mathscr{A}_\eta$.

Proof. Let η_θ be the point of projection of $\boldsymbol{\eta}$ onto $\varkappa(\boldsymbol{\theta})$, i.e. the solution of the equations

$$\left\langle \boldsymbol{\eta}_\theta - \boldsymbol{\eta}(\boldsymbol{\theta}), \frac{\partial \boldsymbol{\eta}}{\partial \theta_i} \right\rangle = 0 ; \quad (i = 1, \ldots, m) ,$$

which satisfies the equality

$$\boldsymbol{\eta}_\theta - \boldsymbol{\eta} = \sum_{j=1}^{m} k_j^{(\theta)} \frac{\partial \boldsymbol{\eta}(\boldsymbol{\theta})}{\partial \theta_j}$$

for some $k_1^{(\theta)}, \ldots, k_m^{(\theta)}$. Using the implicit function theorem (cf. [2], Theorem 211) we may verify that the mapping $\theta \mapsto \boldsymbol{\eta}_\theta$ is continuous in a neighbourhood V_{θ^-} of $\bar{\theta}$.

We have $\|\boldsymbol{\eta} - \boldsymbol{\eta}_{\theta^-}\| = \min \{\|\boldsymbol{\eta} - \mathbf{z}\| : \mathbf{z} \in \varkappa(\bar{\theta})\} < r$, because $\eta(\bar{\theta}) \in \mathscr{A}_\eta$. Hence $\|\boldsymbol{\eta}_{\theta^-} - \boldsymbol{\eta}(\bar{\theta})\| < r$ (Proposition A 2). From the continuity of the mappings $\theta \mapsto \boldsymbol{\eta}(\theta)$, $\theta \mapsto \boldsymbol{\eta}_\theta$ we have a neighbourhood $U_{\theta^-} \subset V_{\theta^-}$ such that

$$\|\boldsymbol{\eta}_\theta - \boldsymbol{\eta}\| < r, \quad \|\boldsymbol{\eta}_\theta - \boldsymbol{\eta}(\theta)\| < r \,; \quad (\theta \in U_{\theta^-}) \,.$$

Therefore, according to (A 13) we have $\boldsymbol{\eta}(\theta) \in \mathscr{A}_\eta$; $(\theta \in U_{\theta^-})$. □

Let us denote

$$t(\mathbf{y}) := \operatorname{Arg}\min_t \|\gamma(t) - \mathbf{y}\| \,, \tag{A 15}$$

$$S_\gamma(t) := \left\{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \frac{\mathrm{d}}{\mathrm{d}t} \|\gamma(t) - \mathbf{y}\|^2 > 0\right\} \tag{A 16}$$

$$Q_\gamma(t) := \left\{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \frac{\mathrm{d}}{\mathrm{d}t} \|\gamma(t) - \boldsymbol{\eta}\|^2_{t(\mathbf{y})} < \frac{\mathrm{d}}{\mathrm{d}t} \|\gamma(t) - \boldsymbol{\eta}\|^2\right\} . \tag{A 17}$$

Proposition A 6. Let γ be a geodesics, $\gamma(\bar{t}) = \boldsymbol{\eta}(\bar{\theta}) \in \mathscr{A}_\eta$, $\boldsymbol{\eta}(U_{\theta^-}) \subset \mathscr{A}_\eta$, $\gamma(t) \in (U_{\theta^-})$, $t < \bar{t}$.

Then

$$W_r \cap [S_\gamma(\bar{t}) - S_\gamma(t)] = W_r \cap [Q_\gamma(\bar{t}) - Q_\gamma(t)] \,. \tag{A 18}$$

Proof. From Proposition A 3 it follows that the function $t \mapsto (\mathrm{d}/\mathrm{d}t) \|\gamma(t) - \boldsymbol{\eta}\|^2$ is increasing as long as $\gamma(t) \in U_{\theta^-}$. Evidently $(\mathrm{d}/\mathrm{d}t) \|\gamma(t) - \mathbf{y}\|^2_{t(\mathbf{y})} = 0$. Hence, for $t < \bar{t}$, $\gamma(t) \in U_{\theta^-}$ we have

$$\mathbf{y} \in Q_\gamma(\bar{t}) - Q_\gamma(t) \Leftrightarrow t \leqq t(\mathbf{y}) < \bar{t} \,. \tag{A 19}$$

The halfspaces $S_\gamma(\bar{t})$, resp. $S_\gamma(t)$, are limited by the hyperplanes $\varkappa_\gamma(\bar{t})$, resp. $\varkappa_\gamma(t)$. Therefore from Proposition A 4 it follows that

$$\langle t, \bar{t}) = \{t(\mathbf{y}) : \mathbf{y} \in W_r \cap [S_\gamma(\bar{t}) - S_\gamma(t)]\} \,.$$

Comparing this with Eq. (A 19) we obtain Eq. (A 18). □

Take a point $\bar{\theta} \in U$ such that $\boldsymbol{\eta}(\bar{\theta}) \in \mathscr{A}_\eta$. In the remaining part of the Appendix we shall introduce adequate local coordinates on $\mathscr{E}$ and local coordinates on $\mathbb{R}^N$ in a neighbourhood of the point $\boldsymbol{\eta}(\bar{\theta})$.

Take m geodesics $\gamma^{(1)}, \ldots, \gamma^{(m)}$ such that

$$\gamma^{(i)}(0) = \boldsymbol{\eta}(\bar{\theta}) \,; \quad (i = 1, \ldots, m) \,. \tag{A 20}$$

$$\dot{\gamma}^{(i)}(0), \dot{\gamma}^{(j)}(0)\rangle = 0 \,; \quad (i \neq j) \,.$$

We introduce new local coordinates on $\mathscr{E}$, $t_1 = \tau_1(\theta), \ldots, t_m = \tau_m(\theta)$ by the equations

(A 21) $$\langle \boldsymbol{\eta}(\boldsymbol{\theta}) - \gamma^{(i)}(t_i), \dot{\gamma}^{(i)}(t_i)\rangle = 0\,; \quad (i = 1, \ldots, m)\,.$$

From the implicit function theorem (cf. [2]) it follows that the functions $\tau_1(\boldsymbol{\theta}), \ldots$ $\ldots, \tau_m(\boldsymbol{\theta})$ are one-to-one and differentiable in a neighbourhood $V_{\bar{\boldsymbol{\theta}}} \subset U_{\bar{\boldsymbol{\theta}}}$.

Analogically, we define new coordinates $x_1 = \xi_1(\mathbf{y}), \ldots, x_N = \xi_N(\mathbf{y})$ in the set $\mathscr{G}_{\bar{\boldsymbol{\theta}}} := \{\mathbf{y} : \mathbf{y} \in \mathbb{R}^N, \mathbf{y} \in \varkappa(\boldsymbol{\theta}), \boldsymbol{\theta} \in V_{\bar{\boldsymbol{\theta}}}\}$ by the equations

(A 22) $$\langle \boldsymbol{\eta}[\boldsymbol{\theta}^*(\mathbf{y})] - \gamma^{(i)}(x_i), \dot{\gamma}^{(i)}(x_i)\rangle = 0\,; \quad (i = 1, \ldots, m)$$
$$\xi_i(\mathbf{y}) := \langle \mathbf{y} - \boldsymbol{\eta}[\boldsymbol{\theta}^*(\mathbf{y})], \mathbf{w}^{(i)}[\boldsymbol{\theta}^*(\mathbf{y})]\rangle\,; \quad (i = m + 1, \ldots, N)\,,$$

where $\mathbf{w}^{(m+1)}(\boldsymbol{\theta}), \ldots, \mathbf{w}^{(N)}(\boldsymbol{\theta})$ is an orthonormal basis of $\{\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}) : \mathbf{y} \in \varkappa(\boldsymbol{\theta})\}$ and $\boldsymbol{\theta}^*(\mathbf{y})$ is the solution of Eq. (4) which is in $V_{\bar{\boldsymbol{\theta}}}$.

Other coordinates $z_1 = \zeta_1(\mathbf{y}), \ldots, z_N = \zeta_N(\mathbf{y})$ can be introduced as follows. First we define $z_1, \ldots, z_m$ by

(A 23) $$\langle \mathbf{y} - \gamma^{(i)}(z_i), \dot{\gamma}^{(i)}(z_i)\rangle = 0\,; \quad (i = 1, \ldots, m).$$

Further we denote by $\tilde{\boldsymbol{\theta}}(\mathbf{y}) \in U$ the vector defined by

$$\{\boldsymbol{\eta}[\tilde{\boldsymbol{\theta}}(\mathbf{y})]\} = \mathscr{A}_{\boldsymbol{\eta}} \cap \bigcap_{i=1}^{m} \varkappa_i[\zeta_i(\mathbf{y})]$$

(cf. Eq. (A 7) for the definition of $\varkappa_i(t) := \varkappa_{\gamma^{(i)}}(t)$). Denote by $\mathbf{r}^{(m+1)}(\mathbf{y}), \ldots, \mathbf{r}^{(N)}(\mathbf{y})$ an orthonormal basis of $\bigcap_{i=1}^{m} \varkappa_i[\zeta_i(\mathbf{y})]$. Let us define

(A 24) $$\zeta_j(\mathbf{y}) := \langle \mathbf{y} - \boldsymbol{\eta}[\tilde{\boldsymbol{\theta}}(\mathbf{y})], \mathbf{r}^{(j)}(\mathbf{y})\rangle\,; \quad (j = m + 1, \ldots, N)\,.$$

Evidently

(A 25) $$\boldsymbol{\theta}^*(\mathbf{y}) = \bar{\boldsymbol{\theta}} \Rightarrow \tilde{\boldsymbol{\theta}}(\mathbf{y}) = \bar{\boldsymbol{\theta}},\ \mathbf{r}^{(j)}(\mathbf{y}) = \mathbf{w}^{(j)}\,[\hat{\boldsymbol{\theta}}(\mathbf{y})] = \mathbf{w}^{(j)}(\bar{\boldsymbol{\theta}})\,,$$
$$x_1 = \ldots = x_m = 0\,, \quad z_1 = \ldots = z_m = 0 \quad x_{m+1} = z_{m+1}, \ldots, x_N = z_N\,.$$

The functions $\zeta_1, \ldots, \zeta_N$ are one-to-one and differentiable in the set $\mathscr{G}_{\bar{\boldsymbol{\theta}}}$ with $V_{\bar{\boldsymbol{\theta}}}$ choosen adequatelly. Evidently

(A 26) $$\xi_i(\mathbf{y}) = \tau_i[\boldsymbol{\theta}^*(\mathbf{y})]\,,$$
$$\zeta_i(\mathbf{y}) = \tau_i[\tilde{\boldsymbol{\theta}}(\mathbf{y})]\,; \quad (i = 1, \ldots, m)$$

We shall compute the Jacobi matrices of the mappings $\boldsymbol{\theta} \mapsto \mathbf{t}$, $\mathbf{x} \mapsto \mathbf{y}$, $\mathbf{y} \mapsto \mathbf{z}$.

Differentiating Eqs. (A 21) with respect to θ_j we obtain

$$\left\langle \frac{\partial \boldsymbol{\eta}(\boldsymbol{\theta})}{\partial \theta_j}, \gamma^{(i)}(t_i)\right\rangle + \sum_{k=1}^{m} \frac{\partial}{\partial t_k} \langle \boldsymbol{\eta}(\boldsymbol{\theta}) - \gamma^{(i)}(t_i), \dot{\gamma}^{(i)}(t_i)\rangle \frac{\partial \tau_k}{\partial \theta_j} = 0\,.$$

Hence

(A 27) $$\left.\frac{\partial \tau_i}{\partial \theta_j}\right|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}} = \left\langle \frac{\partial \boldsymbol{\eta}(\bar{\boldsymbol{\theta}})}{\partial \theta_j}, \dot{\gamma}^{(i)}(0)\right\rangle; \quad (i, j = 1, \ldots, m)\,.$$

Analogously, from Eqs. (A 23) we obtain

$$\text{(A 28)}\quad \left.\frac{\partial \zeta_i}{\partial y_j}\right|_{\theta^*(\mathbf{y})=\bar{\theta}} = \frac{\dot{\gamma}_j^{(i)}(0)}{\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\|\gamma^{(i)}(t)-\mathbf{y}\|^2_{t=0}}\,; \quad (i=1,\ldots,m,\ j=1,\ldots,N)$$

From (A 24) and (A 26) we obtain

$$\zeta_i(\mathbf{y}) = \langle \mathbf{y} - \boldsymbol{\eta}[\tau^{-1}(\zeta_1(\mathbf{y}),\ldots,\zeta_m(\mathbf{y}))]\,,\ \ \mathbf{r}^{(i)}(\mathbf{y})\rangle\,;\quad (i = m+1,\ldots,N)\,.$$

Hence

$$\text{(A 29)}\qquad \left.\frac{\partial \zeta_i}{\partial y_j}\right|_{\hat{\theta}(\mathbf{y})=\bar{\theta}} = r_j^{(i)}(\mathbf{y}) + \sum_{s=1}^{m} Q_{is}\frac{\partial \zeta_s}{\partial y_j} = w_j^{(i)}\,;$$

$$(i = m+1,\ldots,N, j = 1,\ldots,N)$$

where

$$Q_{is} := \frac{\partial}{\partial z_s}\langle \mathbf{y} - \boldsymbol{\eta}[\tau^{-1}(z_1,\ldots,z_m), \mathbf{w}^{(i)}(\bar{\theta})\rangle\big|_{z_1=\ldots=z_m=0} = 0\,.$$

From Eqs. (A 22) and (A 26) we obtain

$$\mathbf{y} = \eta[\tau^{-1}(x_1,\ldots,x_m) + \sum_{j=m+1}^{N} x_j \mathbf{w}^{(j)}[\tau^{-1}(x_1,\ldots,x_m)]\,.$$

It follows that for $i = 1,\ldots,N$

$$\text{(A 30)}\qquad \left.\frac{\partial y_i}{\partial x_j}\right|_{\theta^*(y)=\bar{\theta}} = \left.\frac{\partial \boldsymbol{\eta}_i[\tau^{-1}(x_1,\ldots,x_m)]}{\partial x_j}\right|_{x_1=\ldots=x_m=0}\,;\quad (j=1,\ldots,m)$$

$$= w^i_{(j)}(\bar{\theta})\,;\quad (j = m+1,\ldots,N)\,.$$

Proposition A 7. If $\boldsymbol{\Sigma} = \mathbf{I}$ then

$$\{\mathbf{J}\}_{ij} := \left.\frac{\partial z_i}{\partial x_j}\right|_{x_1=\ldots=x_m=0} \begin{cases} = 1\,; & i=j\,, \\ = 0\,; & i \neq j\,; \quad (i,j=1,\ldots,N)\,. \end{cases}$$

Proof. Let us denote $\mathbf{d}^{(i)} := \dot{\gamma}^{(i)}(0)/\frac{1}{2}(\mathrm{d}^2/\mathrm{d}t^2)\,\|\gamma^{(i)}(t)-\mathbf{y}\|_0^2$.

We have $\{\mathbf{J}\}_{ij} = \sum_k (\partial z_i/\partial y_k)(\partial y_k/\partial x_j)$. Hence from Eqs. (A 28)–(A 30) we obtain

(A 31)

$$\mathbf{J} = \begin{bmatrix} \mathbf{d}^{(1)\prime} \\ \vdots \\ \mathbf{d}^{(m)\prime} \\ \mathbf{w}'_{m+1}(\theta) \\ \vdots \\ \mathbf{w}'_N(\theta) \end{bmatrix} \left(\left.\frac{\partial\boldsymbol{\eta}[\tau^{-1}(\mathbf{x})]}{\partial x_1}\right|_{\mathbf{x}=\mathbf{0}}, \ldots, \left.\frac{\partial\boldsymbol{\eta}[\tau^{-1}(\mathbf{x})]}{\partial x_m}\right|_{\mathbf{x}=\mathbf{0}},\ \ \mathbf{w}_{m+1}(\theta),\ldots,\mathbf{w}_N(\theta)\right) = \mathbf{I}$$

since, if $\theta^*(\mathbf{y}) = \bar{\theta}$ then $\partial\boldsymbol{\eta}[\tau^{-1}(\mathbf{x})]/\partial x_i = \gamma^{(i)\prime}(0)$. □

APPENDIX B (COMPUTATION)

It may be useful to consider the computational aspect when computing the density $q(\hat{\boldsymbol{\theta}} \mid \boldsymbol{\theta})$ and the level of regularity $1-p_0$. To be concrete, let us consider the following example.

Take

$$\eta_x(\boldsymbol{\theta}) = e^{\theta_1 x} \sin \theta_2 x \; ; \quad \theta_1 \in (0, 10)$$
$$\theta_2 \in (0, 2\pi)\,,$$

take $\boldsymbol{\Sigma} = \mathbf{I}$, and take 4 design points $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4$.

The program for the computation of $f(\hat{\boldsymbol{\theta}} \mid \boldsymbol{\theta})$:

Input variables: $\theta_1, \theta_2, \hat{\theta}_1, \hat{\theta}_2$ (4 numbers)

Subroutines:

(A) $$\eta_i(\boldsymbol{\theta}) = \mathrm{e}^{i\theta_1} \sin i\theta_2 \; ; \quad (i = 1, 2, 3, 4)$$

(B) $$\frac{\partial \eta_i(\boldsymbol{\theta})}{\partial \theta_1} = i\, \mathrm{e}^{i\theta_1} \sin i\theta_2$$

(C) $$\frac{\partial \eta_i(\boldsymbol{\theta})}{\partial \theta_2} = i\, \mathrm{e}^{i\theta_1} \cos i\theta_2$$

(D) $$\frac{\partial^2 \eta_i(\boldsymbol{\theta})}{\partial \theta_1^2} = i^2\, \mathrm{e}^{i\theta_1} \sin i\theta_2$$

(E) $$\frac{\partial^2 \eta_i(\boldsymbol{\theta})}{\partial \theta_1\, \partial \theta_2} = i^2\, \mathrm{e}^{i\theta_1} \cos i\theta_2$$

(F) $$\frac{\partial^2 \eta_i(\boldsymbol{\theta})}{\partial \theta_2^2} = -\, i^2\, \mathrm{e}^{i\theta_1} \sin i\theta_2$$

Subroutines for matrices:

(G) $$M_{jk}(\boldsymbol{\theta}) = \sum_{i=1}^{4} \frac{\partial \eta_i(\boldsymbol{\theta})}{\partial \theta_j} \frac{\partial \eta_i(\boldsymbol{\theta})}{\partial \theta_k} \; ; \quad (j, k = 1, 2)$$

(H) $$\mathbf{M}(\boldsymbol{\theta}) \mapsto \mathbf{M}^{-1}(\boldsymbol{\theta})$$

(I) $$P_{jk}^{\boldsymbol{\theta}} = \sum_{p,q=1}^{2} \frac{\partial \eta_j(\boldsymbol{\theta})}{\partial \theta_p} \{M^{-1}(\boldsymbol{\theta})\}_{pq} \frac{\partial \eta_k(\boldsymbol{\theta})}{\partial \theta_q}$$

Use the subroutines (A)–(I) for $\boldsymbol{\theta} = (\hat{\theta}_1, \hat{\theta}_2)$, the subroutine (A) for $\boldsymbol{\theta} = (\theta_1, \theta_2)$ and compute Eq. (9) for different inputs $\theta_1, \theta_2, \hat{\theta}_1, \hat{\theta}_2$.

The program for the computation of $(1 - p_0)$:

The main idea of the algorithm is that through any point $\boldsymbol{\theta} = (\theta_1, \theta_2) \in (0, 10) \times$ $\times\, (0, 2\pi)$ and in any direction given by $\dot{\boldsymbol{\theta}} := \mathrm{d}\boldsymbol{\theta}/\mathrm{d}t$ we can draw a unique geodesics which is a solution if Eqs. (A 3), but for the natural parameter t_{nat}, where $\mathrm{d}t_{nat}/\mathrm{d}t = \|\mathrm{d}\boldsymbol{\eta}/\mathrm{d}t\|$ (cf. Eq. (A 2)).

Input: $\theta_1, \theta_2, \dot\theta_1, \dot\theta_2$ (4 numbers)

Subroutines: (B)–(F)

Subroutine "derivatives":

(J)
$$\frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} = \frac{\partial\boldsymbol{\eta}(\boldsymbol{\theta})}{\partial\theta_1}\dot\theta_1 + \frac{\partial\boldsymbol{\eta}(\boldsymbol{\theta})}{\partial\theta_2}\dot\theta_2$$

(K)
$$\frac{\mathrm{d}^2\boldsymbol{\eta}}{\mathrm{d}t^2}[v, w] = \frac{\partial^2\boldsymbol{\eta}(\boldsymbol{\theta})}{\partial\theta_1^2}(\dot\theta_1)^2 + 2\frac{\partial^2\boldsymbol{\eta}(\boldsymbol{\theta})}{\partial\theta_1\,\partial\theta_2}\dot\theta_1\dot\theta_2 + \frac{\partial^2\boldsymbol{\eta}(\boldsymbol{\theta})}{\partial\theta_2^2}(\dot\theta_2)^2 + \frac{\partial\boldsymbol{\eta}(\boldsymbol{\theta})}{\partial\theta_1}v + \frac{\partial\boldsymbol{\eta}(\boldsymbol{\theta})}{\partial\theta_2}w$$

where v, w are unknown input variables (interpreted as $v = \ddot\theta_1$, $w = \ddot\theta_2$),

(L)
$$\frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t_{nat}} = \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t}\Big/\left\|\frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t}\right\|$$

(M)
$$\frac{\mathrm{d}^2\boldsymbol{\eta}[v,w]}{\mathrm{d}t_{nat}^2} = \frac{\dfrac{\mathrm{d}^2\boldsymbol{\eta}}{\mathrm{d}t^2}[v, w] - \dfrac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t}\left(\dfrac{\mathrm{d}}{\mathrm{d}t}\left\|\dfrac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t}\right\|\right)\Big/\left\|\dfrac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t}\right\|}{\left\|\dfrac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t}\right\|^2}$$

Linear equations: Compute v, w as the solution of the linear equations

(EQ)
$$\sum_{i=1}^{4}\frac{\mathrm{d}^2\eta_i[v,w]}{\mathrm{d}t_{nat}^2}\frac{\partial\eta_i(\boldsymbol{\theta})}{\partial\theta_1} = 0$$
$$\sum_{i=1}^{4}\frac{\mathrm{d}^2\eta_i[v,w]}{\mathrm{d}t_{nat}^2}\frac{\partial\eta_i(\boldsymbol{\theta})}{\partial\theta_2} = 0$$

(cf. Eqs. (A 3))

Put v, w into (K), (M) and compute

$$\varrho(\theta_1, \theta_2, \dot\theta_1, \dot\theta_2) = \left\|\frac{\mathrm{d}^2\boldsymbol{\eta}[v,w]}{\mathrm{d}t_{nat}^2}\right\|^{-1}.$$

For different inputs $\theta_1, \theta_2, \dot\theta_1, \dot\theta_2$ compute

$$r = \min\{\varrho(\theta_1, \theta_2, \dot\theta_1, \dot\theta_2) : \theta_1 \in (0, 10), \theta_2 \in (0, 2\pi),$$
$$\dot\theta_1 \in \langle 0, 1\rangle, \dot\theta_2 \in \langle 0, 1\rangle, \dot\theta_1^2 + \dot\theta_2^2 = 1\}.$$

Compute p_0 from

$$\chi_4^2(p_0) = r^2$$

where $\chi_4^2(p_0)$ is the $(1 - p_0)$ quantile of the χ^2 p.d. with 4 degrees of freedom.

(Received July 18, 1983.)

REFERENCES

[1] B. Hostinský: Diferenciální geometrie křivek a ploch (Differential geometry of curves and surfaces). Přírodovědecké nakladatelství JČSMF, Praha 1950.

[2] V. Jarník: Diferenciální počet II (Differential calculus). NČSAV, Praha 1956.

[3] J. Milnor: Morse Theory. Princeton Univ. Press, Princeton, N. J. 1963.

[4] A. Pázman: Geometry of gaussian nonlinear regression — parallel curves and confidence intervals. Kybernetika *18* (1982), 376—396.

[5] A. Pázman: Nonlinear least squares — uniqueness versus ambiguity. Math. Operationsforsch. Statist. Ser. Statistics *14* (1984) (to appear.)

[6] A. Rényi: Teorie pravděpodobnosti (Probability theory). Academia, Praha 1972.

[7] S. Sternberg: Lectures on Differential Geometry. Second printing. Prentice-Hall, Englewood Cliffs 1965.

RNDr. Andrej Pázman, DrSc., Matematický ústav SAV (Mathematical Institute — Slovak Academy of Sciences), ul. Obrancov mieru 49, 814 73 Bratislava. Czechoslovakia.