View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

SFB 649 Discussion Paper 2006-080

The Uniqueness of Extremum Estimation

Volker Krätschmer*

SFB 649 ECONOMIC RISK BERLIN

* Institute of Mathematics, Berlin University of Technology, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

The uniqueness of extremum estimation

Volker Krätschmer *

Institute of Mathematics, Berlin University of Technology,

Germany

Abstract

Let W denote a family of probability distributions with parameter space Γ, and W_G be a subfamily of W depending on a mapping $G : \Theta \to \Gamma$. Extremum estimations of the parameter vector $\vartheta \in \Theta$ are considered. Some sufficient conditions are presented to ensure the uniqueness with probability one. As important applications, the maximum likelihood estimation in curved exponential families and nonlinear regression models with independent disturbances as well as the maximum likelihood estimation of the location and scale parameters of Gumbel distributions are treated.

Keywords: Extremum estimation, Sard's theorem, nonlinear regression, curved exponential families, Gumbel distributions.

AMS classification 62F10, 62F11

JEL classification C13, C16

1 Introduction

Extremum estimation designates a principle often used for estimating unknown parameters: Starting from data those parameter vectors are selected which maximize or minimize a certain function defined over

*This research was supported by Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

the parameter space. Some estimation methods currently applied in practical situations of estimation are subsumed under the principle of extremum estimation, for example the maximum likelihood and least squares estimation. Employing the extremum estimation in practice causes the problem to choose some parameter vector if the related maximation (or minimation) problem has more than one solution. Under general conditions the question of the uniqueness of extremum estimation seems to be still open. A known result of some special cases was presented by Pazman (cf. [8]). He showed that in some curved exponential families the maximum likelihood method has at most one solution with probability one.

The present paper is an attempt to generalize the result of Pazman to a greater variety of applications, with respect to the extremum estimation.

Since there is no essential difference between maximation and minimization, we shall restrict ourselves to estimation that maximizes a certain function which is dependent on the parameters. After introducing some general notations we shall present in section 3 our main result. We shall deal with the estimation of parameter vectors of distributions which are dominated by Lebesgue measures. The fundamental idea is to employ Sard's theorem to ensure that almost surely the estimation has at most one solution, adapting the line of reasoning that Pazman followed in the above mentioned paper. It may be simplified when considering uniqueness of nonlinear least squares estimation. So the reader is referred to Pazman's monograph on nonlinear regression (cf. [9]) to get a first idea.

The main result of the paper relies on the assumption that for almost every observation from the sample space the respective objective function has only non degenerated critical points. In section 4 this assumption will be replaced by some conditions to restate the main result. The obtained criteria to ensure uniqueness of extremum estimation almost surely will be applied to the maximum likelihood estimation in nonlinear regression models with independent disturbances (section 5) as well as curved exponential families (section 6), and also to maximum likelihood estimation of the location and scale parameter of a Gumbel distribution (section 7). Appendix A provides a crucial auxiliary result, whereas appendix B deals with the proof of the main theorem. Finally appendix C reviews Lindelöf's theorem and useful special versions of Sard's theorem.

2 Notations and preliminaries

For a differentiable mapping $f : H \to \mathbf{R}^t$ defined on an open subset H of $\mathbb{R}^s$, the Jacobian of f at x will be denoted by $J_x(f)$, its rank will be symbolized by $rank(J_x(f))$. As usual f will be called a C^r-mapping ($r \in \mathbb{N}$) if it is r times continuously differentiable. f is said to be a C^r-immersion/C^r-submersion ($r \in \mathbb{N}$) if it is a C^r-mapping with $rank(J_x(f)) = s/rank(J_x(f)) = t$ for all $x \in H$. If $t = n$, and if f is an homeomorphism from H onto the open subset $f(H)$ such that $f : H \to f(H)$ and its inverse $f^{-1} : f(H) \to H$ are C^r-immersions, then f is called a C^r-diffeomorphism onto $f(H)$. The Jacobians of a C^r-diffeomorphism are nonsingular. In the case of $t = 1$ $\nabla_x f$ stands for the gradient of f at x, whereas $\nabla\nabla_x f$ will be used for the Hessian of f at x if f is twice differentiable. If $\nabla_x f = 0$, then we shall speak of x as a critical point, which will be called degenerated provided that $\nabla\nabla_x f$ is singular.

For $m \in \mathbb{N}_0$ and $r \in \mathbb{N}$ a subset $A \subseteq \mathbb{R}^s$ is called a $m-$dimensional C^r-submanifold of $\mathbb{R}^s$ if for any $x \in A$ there exists a C^r-diffeomorphism f from an open neighbourhood U of x in $\mathbb{R}^s$ onto the open subset $f(U)$ of $\mathbb{R}^s$ such that $f(U \cap A) = f(U) \cap \{(x_1, ..., x_s) \in \mathbb{R}^s \mid x_{m+1} = ... = x_s = 0\}$ (cf. [1], Definition 2.1.1). Note that in view of Lindelöf's theorem (cf. appendix C, Proposition C.2) a $0-$dimensional C^r-submanifold of $\mathbb{R}^s$ is an at most countable set. If $g : U \to \mathbb{R}^t$ denotes a C^r-submersion from an open subset U of $\mathbb{R}^s$ into $\mathbb{R}^t$, then we know by the regular value theorem (cf. [4], p.14 or p.22, Theorem 3.2) that for each $y \in \mathbb{R}^t$ the fibre $g^{-1}(\{y\})$ is empty or a $(s-t)-$dimensional C^r-submanifold of $\mathbb{R}^s$.

The transpose of a vector $\underline{x}$ in a standard euclidean space or a matrix M will be indicated by $\underline{x}'$ and M' respectively.

Furthermore the notation λ^t will be employed for the Lebesgue measure on $\mathbb{R}^t$.

Throughout the paper we shall consider the following setting of extremum estimation: Let Y be a random vector with measurable space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, where $\mathcal{X}$ is an open subset of $\mathbb{R}^n$, and $\mathcal{B}(\mathcal{X})$ denotes the $\sigma-$algebra of Borel subsets of $\mathcal{X}$. We suppose that the distribution Q_Y of Y may be parameterized by a mapping $G : \Theta \to \mathbb{R}^m$, i.e. it belongs to a familiy $W_G = \{P_{G(\vartheta)} \mid \vartheta \in \Theta\}$ of probability distributions, which in addition we assume to be dominated by $\lambda^n \mid Borel(\mathcal{X})$. Of course the parameter ϑ should be identified, which implies

(2.1) G is injective.

Starting from a realisation y of the random vector Y, we are interested in estimating the (unknown) vector ϑ by maximizing the mapping $l(y;\cdot)\circ G =: l_G(y;\cdot)$ defined by a given mapping $l(\cdot;\cdot) : \mathcal{X}\times\Gamma \to \mathbf{R}$, where Γ stands for an open subset of $\mathbb{R}^m$ enclosing $G(\Theta)$. Additionally, it will be assumed

(2.2) $G(\Theta) = \bigcup_{j=1}^{\infty} G^j(\Theta^j)$, where G^j denotes a C^2-mapping defined on an open subset Θ^j of $\mathbf{R}^{r_j}$,

(2.3) $l(\cdot;\cdot)$ is a C^2-mapping.

3 Statement of the main result

The main result relies on the following additional assumptions.

(3.1) Under (2.2), (2.3) the set L_j, consisting of all $y \in \mathcal{X}$ such that $l(y,\cdot)\circ G^j$ has degenerated critical points, is a λ^n-null set for each $j \in \mathbb{N}$.

(3.2) Under (2.2), (2.3) every set L_{ij} $(i,j \in \mathbb{N})$ of all $y \in \mathcal{X}$ with $\nabla_{\vartheta^i} l(y,\cdot)\circ G^i = 0, \nabla_{\vartheta^j} l(y,\cdot)\circ G^j = 0$ and $\nabla_y l(\cdot, G^i(\vartheta^i)) = \nabla_y l(\cdot, G^j(\vartheta^j))$ for some $(\vartheta^i,\vartheta^j) \in \Theta^i\times\Theta^j, G^i(\vartheta^i) \neq G^j(\vartheta^j)$, is a λ^n-null set.

Later on we shall replace (3.1) by some sufficient conditions.

Theorem 3.1 *Let the assumptions (2.1) - (2.3) be fulfilled. If the mappings G^j $(j \in \mathbb{N})$ are C^2- immersions, and if the conditions (3.1), (3.2) are valid, then*

$$K := \{y \in \mathcal{X} \mid l_G(y;\vartheta) = l_G(y;\tilde{\vartheta}) = \max_{\vartheta\in\Theta} l_G(y;\vartheta) \textit{ for some different } \vartheta,\tilde{\vartheta} \in \Theta\} \in \mathcal{B}(\mathcal{X})$$

with $\mathrm{Q}_Y(K) = 0$.

The proof is delegated to appendix B.

In the following we try to substitute condition (3.1). Our main tool will be provided by Lemma A.1 (cf. appendix A). Unfortunately, it can be applied directly only in the case that the dimension n of the sample space $\mathcal{X}$ coincides with the dimension m of the parameter space Γ. But if we are in the position to modify the maximization problem in a suitable way we would be able to draw on Lemma A.1. We shall treat the cases $n \geq m$ and $m \geq n$ separately.

4 Specializations

In the case of $n \geq m$ we can use Lemma A.1 if we may reduce the dimension of the sample space by the following condition:

(4.1) There exist a C^2-submersion $T : \mathcal{X} \to \mathbf{R}^m$ and a real-valued C^2-mapping l^* on $T(\mathcal{X}) \times \Gamma$ such that $l(y; \cdot) = l^*(T(y); \cdot)$ holds for arbitrary $y \in \mathcal{X}$.

Since T is a C^1-submersion, $T(\mathcal{X})$ is an open subset of $\mathbb{R}^m$, and moreover, $\lambda^n(T^{-1}(B)) = 0$ if $\lambda^m(B) = 0$ (cf. [8], Proposition A3). Then a direct application of Theorem 3.1 with Lemma A.1 leads to the following result concerning the uniqueness of the extremum estimation of ϑ.

Theorem 4.1 *Let the condition (4.1) as well as the assumptions (2.1) - (2.3) be valid, let the mappings G^j $(j \in \mathbb{N})$ be C^2-immersions, and let condition (3.2) be fulfilled. If there exist some C^1-diffeomorphism g from an open subset U of $\mathbb{R}^m$ onto an open subset $g(U)$ of $\mathbb{R}^m$ and a C^1-mapping $\phi : \Gamma \to \mathbb{R}^m$ such that $t - \phi(\gamma) \in U$ and $\nabla_\gamma l^*(t; \cdot) = g(t - \phi(\gamma))$ hold for $(t, \gamma) \in T(\mathcal{X}) \times \Gamma$, then*

$$K := \{y \in \mathcal{X} \mid l_G(y; \vartheta) = l_G(y; \tilde{\vartheta}) = \max_{\vartheta \in \Theta} l_G(y; \vartheta) \textit{ for some different } \vartheta, \tilde{\vartheta} \in \Theta\} \in \mathcal{B}(\mathcal{X})$$

with $\mathrm{Q}_Y(K) = 0$.

Remark 4.2 *If $n = m$, condition (4.1) is satisfied, choosing for T the restriction of the identity mapping on $\mathbb{R}^n$ to $\mathcal{X}$. Therefore we may restate in this case Theorem 4.1 without condition (4.1).*

Now let $m \geq n$. For reduction of the dimension of Γ we want to assume:

(4.2) Under (2.2), (2.3) there exist an open subset H of $\mathcal{X} \times \mathbb{R}^n$ with nonvoid $H_y := \{x \in \mathbb{R}^n \mid (y, x) \in H\}$ $(y \in \mathcal{X})$, a C^2-mapping $l^* : H \to \mathbb{R}$ and a sequence $(F^j)_j$ of mappings $F^j : A^j \to \mathbb{R}^n$ such that the following properties are satisfied

a) F^j is a C^2-immersion on an open subset of $\mathbb{R}^{p_j}$ for every $j \in \mathbb{N}$.

b) If for $y \in \mathcal{X}$ and arbitrary $j \in \mathbb{N}$ the mapping $l(y,\cdot) \circ G^j$ has a degenerated critical point, then the mapping $l^*(y,\cdot) \circ F^j|(F^j)^{-1}(H_y)$ has a degenerated critical point too.

Under condition (4.2) the combination of Theorem 3.1 and Lemma A.1 (cf. appendix A) reads as follows.

Theorem 4.3 *Let the condition (4.2) as well as the assumptions (2.1) - (2.3) be valid, let the mappings* G^j $(j \in \mathbb{N})$ *be* C^2-*immersions, and let condition (3.2) be fulfilled. If there exist some* C^1-*diffeomorphism* g *from an open subset* U *of* $\mathbb{R}^n$ *onto an open subset* $g(U)$ *of* $\mathbb{R}^n$*, an open subset* V *of* $\mathbb{R}^n$ *with* $V \supseteq \bigcup_{y \in \mathcal{X}} H_y$*, and a* C^1-*mapping* $\phi : V \to \mathbb{R}^n$ *such that* $y - \phi(\gamma) \in U$ *and* $\nabla_\gamma l^*(y;\cdot) = g(y - \phi(\gamma))$ *hold for* $(y,\gamma) \in H$*, then*

$$K := \{y \in \mathcal{X} \mid l_G(y;\vartheta) = l_G(y;\tilde{\vartheta}) = \max_{\vartheta \in \Theta} l_G(y;\vartheta) \textit{ for some different } \vartheta, \tilde{\vartheta} \in \Theta\} \in \mathcal{B}(\mathcal{X})$$

with $\mathrm{Q}_Y(K) = 0$.

Proof: Due to Lindelöf's theorem (cf. appendix C, Proposition C.2) we may find a sequence $(U_l \times V_l)_l$ of open subsets of $\mathbb{R}^n \times \mathbb{R}^n$ with $\bigcup_{l=1}^{\infty} U_l \times V_l = H$. In view of Lemma A.1 (cf. appendix A) the set B_{lj} of all $y \in U_l$ with $l^*(y,\cdot) \circ F^j|(F^j)^{-1}(V_l)$ having a degenerated critical point is a λ^n-null set for $l, j \in \mathbb{N}$. Then, condition (4.2) implies condition (3.1), and the statement of Theorem 4.1 follows from Theorem 3.1. ∎

5 Maximum likelihood estimation in nonlinear regression models

Let $Y = F(\alpha) + U$ be a nonlinear regression model with nonstochastic regressors, where

(6.1) Y denotes the random vector of the endogenous variables,

(6.2) $F : A \to \mathbf{R}^n$ stands for the regression function defined on a subset A of $\mathbf{R}^r, r \leq n$, (6.3) $U =: (U_1, ..., U_n)'$ symbolizes the random vector of the disturbances U_i which are supposed to be independently and normally distributed with

$$EU = 0, (Var(U_1), ..., Var(U_n)) =: \tilde{F}(\beta), \quad \tilde{F} : (]0, \infty[)^q \to \mathbf{R}^n, \ \beta \mapsto (\beta_{i_1}, ..., \beta_{i_n})$$

($i_1, ..., i_n$ fixed values with $\{i_1, ..., i_n\} = \{1, ..., q\}$).

It is easy to see that the parameter vector (α, β) is identified if and only if

(6.4) the injectivity of the regression function F is supposed.

Moreover, we assume that

(6.5) a sequence $(p_j)_j$ of positive integers $p_j \leq n$, a sequence $(A^j)_j$ of open subsets A^j of $\mathbf{R}^{p_j}$ and a sequence $(F^j)_j$ of mappings $F^j : A^j \to \mathbf{R}^n$ exist with

$$F(A) = \bigcup_{j=1}^{\infty} F^j(A^j).$$

W denotes the family of normal distributions with independent marginal distributions. $\Gamma := \mathbf{R}^n \times (]0, \infty[)^n$ is a parameter space of W. We can now apply our results of extremum estimation to the maximum likelihood estimation of (α, β).

Theorem 5.1 *Let us retake assumptions (6.1)-(6.5), let $l(y, \cdot)$ denote the log likelihood function w.r.t. W and the realization y of Y. Furthermore let $G : A \times (]0, \infty[)^q \to \Gamma$ be defined by $G(\alpha, \beta) := (F(\alpha), \tilde{F}(\beta))$. If the mappings F^j are C^2-immersions, then*

$$K := \{y \in \mathbb{R}^n \mid l_G(y; \vartheta) = l_G(y; \tilde{\vartheta}) = \max_{\vartheta \in \Theta} l_G(y; \vartheta) \text{ for some different } \vartheta, \tilde{\vartheta} \in \Theta\} \text{ is a Borel subset of } \mathbb{R}^n$$

with $\mathrm{Q}_Y(K) = 0$.

Proof:

Let for a positive integer j the mapping $G^j : A^j \times (]0, \infty[)^q \to \Gamma$ be defined by $G^j(\vartheta^j, \beta) = (F^j(\vartheta^j), \tilde{F}(\beta))$. Then by assumption G fulfills the conditions (2.1), (2.2).

Next let us introduce the open subset $H := \{(y,\mu) \in \mathbb{R}^n \times \mathbb{R}^n \mid \sum\limits_{t\in T_i}(y_t - \mu_t)^2 \neq 0$ for all $i \in \{1,...,q\}\}$ of $\mathbb{R}^n \times \mathbb{R}^n$, where $T_i := \{t \in \{1,...,n\} \mid \tilde{F}_t(\beta) = \beta_i\}$ for $i \in \{1,...,q\}$. Furthermore let us consider the mappings $h : H \to (]0,\infty[)^q$, defined by $h(y,\mu) = (\frac{1}{\sharp T_1}\sum\limits_{t\in T_1}[y_t - \mu_t]^2, ..., \frac{1}{\sharp T_q}\sum\limits_{t\in T_q}[y_t - \mu_t]^2)$, and the mapping $l^* : H \to \mathbb{R}$, defined by

$$l^*(y,\mu) := l(y,\mu,\tilde{F}\circ h(y,\mu)) = -\frac{1}{2}\Bigg(n(\ln(2\pi)+1) + \sum_{i=1}^{q}\sharp T_i[\ln(\sum_{t\in T_i}(y_t-\mu_t)^2) - \ln(\sharp T_i)]\Bigg).$$

By routine procedures it can be shown that l^* and $(F^j)_j$ satisfy condition (4.2). Note that we have

$$\nabla_\beta l(y;\mu,\cdot)\circ\tilde{F} = 0 \Leftrightarrow \beta = (\frac{1}{\sharp T_1}\sum_{t\in T_1}[y_t-\mu_t]^2, ..., \frac{1}{\sharp T_q}\sum_{t\in T_q}[y_t-\mu_t]^2),\ \text{for all } (y,\mu)\in H.$$

Moreover we can define on $U := \{x \in \mathbb{R}^n \mid \sum\limits_{t\in T_i} x_t^2 \neq 0$ for all $i \in \{1,...,q\}\}$ the mapping $g : U \to \mathbf{R}^n$ by $g(x) = \sum\limits_{i=1}^{q}\sum\limits_{t\in T_i}\frac{\sharp T_i x_t}{\sum\limits_{s\in T_i} x_s^2}e_t$, where $\{e_1,...,e_n\}$ denotes the standard basis of $\mathbb{R}^n$. We get for arbitrary $(y,\mu)\in H$

$$y-\mu\ \in U,\ \nabla_\mu l^*(y;\cdot) = g(y-\mu),\ \nabla_y l^*(\cdot;\mu) = -g(y-\mu),$$

g is a bijective indefinitely differentiable function from U onto U with $g^{-1} = g$. Thus $\nabla_y l^*(\cdot;\mu) \neq \nabla_y l^*(\cdot;\tilde{\mu})$ for arbitrary $\mu \neq \tilde{\mu}$, which means that condition (3.2) is fulfilled. The statement of the theorem follows from Theorem 4.3. ■

6 Maximum likelihood estimation in curved exponential families

Let ν be a σ-finite measure on the $\sigma-$algebra $\mathcal{B}(\mathbb{R}^n)$ of Borel subsets of $\mathbb{R}^n$ which is dominated by the Lebesgue measure on $\mathbb{R}^n$. Additionally, let $\mathcal{N} := \{\gamma \in \mathbb{R}^n \mid 0 < \int_{\mathbb{R}^n} exp(\gamma' y)\nu(dy) < \infty\}$ and Γ be an open subset of $\mathbb{R}^n$ enclosed in $\mathcal{N}$. Introducing $\psi : \mathcal{N} \to [-\infty,\infty], \gamma \mapsto \ln(\int_{\mathbb{R}^n} exp(\gamma' y)\nu(dy))$ we consider a minimal exponential family $W := \{P_\gamma \mid \gamma \in \Gamma\}$ of probability distributions having densities

$$f_\gamma : \mathbb{R}^n \to \mathbb{R},\ y \mapsto exp(\gamma' y - \psi(\gamma))\ (\gamma \in \Gamma)$$

with respect to ν. The parameter vector γ is identified since W is a minimal exponential family with parameter space $\Gamma \subseteq \mathcal{N}$ (cf. [2],Theorem 1.13). $G : \Theta \to \Gamma$ denotes an injective mapping which induces the subfamily $W_G := \{P_{G(\vartheta)} \in W \mid \vartheta \in \Theta\}$ of W. Now we may consult Theorem 4.1 concerning the uniqueness of the maximum likelihood estimation of the parameter vector ϑ.

Theorem 6.1 *Let the mapping G satisfy the conditions (2.1), (2.2), let Y be a random vector with distribution $Q_Y \in W_G$, and let $l(y;\cdot) : \Gamma \to \mathbb{R}$ denote the log likelihood function with respect to W and $y \in \mathbb{R}^n$. If the mappings G^j $(j \in \mathbb{N})$ are C^2-immersions, then*

$$K := \{y \in \mathbb{R}^n \mid l_G(y;\vartheta) = l_G(y;\tilde{\vartheta}) = \max_{\vartheta\in\Theta} l_G(y;\vartheta) \text{ for some different } \vartheta, \tilde{\vartheta} \in \Theta\} \text{ is a Borel subset of } \mathbb{R}^n$$

with $\mathrm{Q}_Y(K) = 0$.

Proof:

$\psi|\Gamma$ is indefinitely differentiable (cf. [11], Satz 1.164), which implies that $\phi : \Gamma \to \mathbb{R}^n, \gamma \mapsto \nabla_\gamma\psi$, and $l(\cdot;\cdot) \mid \mathcal{X} \times \Gamma : \mathcal{X} \times int(\Gamma) \to \mathbb{R}$, defined by $l(y;\gamma) = \gamma' y - \psi(\gamma)$, are indefinitely differentiable. Moreover, for arbitrary $y \in \mathcal{X}$ and different $\gamma, \tilde{\gamma} \in \Gamma$ we may observe $\nabla_y l(\cdot;\gamma) = \gamma \neq \tilde{\gamma} = \nabla_y l(\cdot;\tilde{\gamma})$, and furthermore $\nabla_\gamma l(y;\cdot) = y - \nabla_\gamma\psi$. Thus the statement is a direct consequence of Theorem 4.1. ■

Remark:

Theorem 6.1 retains the result of Pazman (cf. [8], Theorem) mentioned in the introduction, which in turn also encompasses Pazman's criterion concerning the uniqueness of nonlinear least squares estimation in nonlinear regression models with nonstochastic regressors and i.i.d. disturbances (cf. [7], Theorem 3 and [9], Corollary 4.4.6).

7 Maximum likelihood estimation of location and scale parameter of a Gumbel distribution

Gumbel distributions build a subfamily of the so called extreme value distributions which play an important role in extreme value theory. The Gumbel distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter

$\sigma > 0$ is defined by the Lebesgue density

$$f_{\mu,\sigma} : \mathbb{R} \to \mathbb{R},\ z \mapsto \frac{1}{\sigma}\exp(-\frac{z-\mu}{\sigma} - \exp(-\frac{z-\mu}{\sigma}))$$

(cf. [5], p. 76). We want to investigate the uniqueness of the maximum likelihood estimation of (μ, σ) based on a random sample $Y := (Y_1, ..., Y_n)$ $(n \geq 2)$ from a Gumbel distribution.

For this purposes let us introduce the mapping

$$l(\cdot;\cdot) : \mathbb{R}^n \times (\mathbb{R}\times]0,\infty[) \to \mathbb{R},\ ((y_1, ..., y_n);(\gamma_1,\gamma_2)) \mapsto n\ln(\gamma_2) + \sum_{i=1}^{n} \ln \circ f_{(0,1)}(\gamma_2 y_i - \gamma_1)$$

Since $\phi : \mathbb{R}\times]0,\infty[\to \mathbb{R}\times]0,\infty[,\ (\mu,\sigma) \mapsto (\frac{\mu}{\sigma}, \frac{1}{\sigma})$ is bijective, we may check easily that for every realization $(y_1, .., y_n)$ of $(Y_1, ..., Y_1)$ the maximum likelihood estimation of (μ, σ) has at most one solution if and only if the mapping $l((y_1, ..., y_n);\cdot)$ has at most one maximizing point. Therefore we might try to apply Theorem 3.1 to $l(\cdot;\cdot)$ and the identity mapping G on $\mathbb{R}\times]0,\infty[$. Obviously, both mappings together satisfy the assumptions (2.1) - (2.3), and G is a C^2-immersion.

Observing that $\frac{df_{(0,1)}}{dz}(z) = f_{(0,1)}(z)(\exp(-z)-1)$ holds for $z \in \mathbb{R}$, routine calculations lead to the following partial derivatives of first and second order for $l((y_1, ..., y_n);\cdot) \circ G$ $((y_1, ..., y_n) \in \mathbb{R}^n)$

$$\frac{\partial l((y_1, ..., y_n);\cdot) \circ G}{\partial \gamma_1}(\gamma_1,\gamma_2) = n - \exp(\gamma_1)\sum_{i=1}^{n} \exp(-\gamma_2 y_i)$$

$$\frac{\partial l((y_1, ..., y_n);\cdot) \circ G}{\partial \gamma_2}(\gamma_1,\gamma_2) = \frac{n}{\gamma_2} - \sum_{i=1}^{n} y_i + \exp(\gamma_1)\sum_{i=1}^{n} y_i \exp(-\gamma_2 y_i)$$

$$\frac{\partial^2 l((y_1, ..., y_n);\cdot) \circ G}{\partial \gamma_1^2}(\gamma_1,\gamma_2) = -\exp(\gamma_1)\sum_{i=1}^{n} \exp(-\gamma_2 y_i)$$

$$\frac{\partial^2 l((y_1, ..., y_n);\cdot) \circ G}{\partial \gamma_1 \partial \gamma_2}(\gamma_1,\gamma_2) = \exp(\gamma_1)\sum_{i=1}^{n} y_i \exp(-\gamma_2 y_i)$$

$$\frac{\partial^2 l((y_1, ..., y_n);\cdot) \circ G}{\partial \gamma_2^2}(\gamma_1,\gamma_2) = -\frac{n}{\gamma_2^2} - \exp(\gamma_1)\sum_{i=1}^{n} y_i^2 \exp(-\gamma_2 y_i)$$

Now let $(\hat{\gamma_1}, \hat{\gamma_2})$ be a critical point of the function $l((y_1, ..., y_n;\cdot) \circ G$. Then $\exp(\hat{\gamma_1}) = \frac{n}{\sum_{i=1}^{n} \exp(-\hat{\gamma}_2 y_i)}$ as well as $\frac{n}{\hat{\gamma}_2} = \sum_{i=1}^n y_i - \exp(\hat{\gamma}_1) \sum_{i=1}^n y_i \exp(-\hat{\gamma}_2 y_i)$, and hence

$$\det \nabla\nabla_{(\hat{\gamma}_1,\hat{\gamma}_2)} l((y_1,...,y_n);\cdot) \circ G = \frac{\sum_{i=1}^n \exp(-\hat{\gamma}_2 y_i)\big(ny_i - \sum_{j=1}^n y_j\big)^2}{\sum_{i=1}^n \exp(-\hat{\gamma}_2 y_i)}$$

Thus the set L of all $(y_1,...,y_n) \in \mathbb{R}^n$ with $l((y_1,...,y_n);\cdot) \circ G$ having a degenerated critical point coincides with the subvector space of all $(y_1,...,y_n) \in \mathbb{R}^n$ with $ny_i = \sum_{j=1}^n y_j$ for $i = 1,...,n$. In particular L is a λ^n-null set.

Let us now consider for $(y_1,...,y_n) \in \mathbb{R}^n$ different critical points $(\hat{\gamma}_1,\hat{\gamma}_2)$ and $(\tilde{\gamma}_1,\hat{\gamma}_2)$ of $l((y_1,...,y_1);\cdot) \circ G$ satisfying $\nabla_{(y_1,...,y_n)} l(\cdot;(\hat{\gamma}_1,\hat{\gamma}_2)) = \nabla_{(y_1,...,y_n)} l(\cdot;(\tilde{\gamma}_1,\tilde{\gamma}_2))$. This means that for each $i \in \{1,...,n\}$ the equation $\hat{\gamma}_2(\exp(\hat{\gamma}_1)\exp(-\hat{\gamma}_2 y_i) - 1) = \tilde{\gamma}_2(\exp(\tilde{\gamma}_1)\exp(-\tilde{\gamma}_2 y_i) - 1)$ is valid. Hence $\hat{\gamma}_2 \neq \tilde{\gamma}_2$, say $\hat{\gamma}_2 > \tilde{\gamma}_2$, and additionally $\hat{\gamma}_2 \exp(\hat{\gamma}_1) = (\hat{\gamma}_2 - \tilde{\gamma}_2)\exp(\hat{\gamma}_2 y_i) + \tilde{\gamma}_2 \exp(\tilde{\gamma}_1)\exp((\hat{\gamma}_2 - \tilde{\gamma}_2)y_i)$ for $i = 1,...,n$, which in turn implies $(\hat{\gamma}_2 - \tilde{\gamma}_2)\exp(\hat{\gamma}_2 y_i) + \tilde{\gamma}_2 \exp(\tilde{\gamma}_1)\exp((\hat{\gamma}_2 - \tilde{\gamma}_2)y_i) = (\hat{\gamma}_2 - \tilde{\gamma}_2)\exp(\hat{\gamma}_2 y_1) + \tilde{\gamma}_2 \exp(\tilde{\gamma}_1)\exp((\hat{\gamma}_2 - \tilde{\gamma}_2)y_1)$ for $i = 1,...,n$.

The mapping $g : \mathbb{R} \to \mathbb{R},\ x \mapsto (\hat{\gamma}_2 - \tilde{\gamma}_2)\exp(\hat{\gamma}_2 x) + \tilde{\gamma}_2 \exp(\tilde{\gamma}_1)\exp((\hat{\gamma}_2 - \tilde{\gamma}_2)x)$, is differentiable, and its derivatives satisfy

$$\frac{dg}{dx}(x) = (\hat{\gamma}_2 - \tilde{\gamma}_2)[\hat{\gamma}_2 \exp(\hat{\gamma}_2 x) + \tilde{\gamma}_2 \exp(\tilde{\gamma}_1)\exp((\hat{\gamma}_2 - \tilde{\gamma}_2)x)] > 0 \ (x \in \mathbb{R})$$

Therefore g is injective, and we may conclude $y_1 = ... = y_n$. Then we know that the set $\hat{L}$ of all $(y_1,...,y_n)$ such that $\nabla_{(y_1,...,y_n)} l(\cdot;(\hat{\gamma}_1,\hat{\gamma}_2)) = \nabla_{(y_1,...,y_n)} l(\cdot;(\tilde{\gamma}_1,\tilde{\gamma}_2))$ holds for some couple $(\hat{\gamma}_1,\hat{\gamma}_2)$ and $(\tilde{\gamma}_1,\hat{\gamma}_2)$ of different critical points of $l((y_1,...,y_1);\cdot) \circ G$ is enclosed in the subvector space of all $(y_1,...,y_n) \in \mathbb{R}^n$ with $y_1 = ... = y_n$. In particular $\hat{L}$ is a λ^n-null set.

Altogether we have shown that $l(\cdot;\cdot)$ and G fulfill the conditions (3.1), (3.2), and this yields the following result concerning the uniqueness of the maximum likelihood estimation of (μ,σ) due to Theorem 3.1.

Theorem 7.1 *Let $Y := (Y_1,...,Y_n)$ $(n \geq 2)$ be a random sample from the Gumbel distribution with the location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$, and let $L(y_1,...,y_n;\cdot)$ denote the likelihood function* *w.r.t. the family of Gumbel distributions and the realization* $(y_1, ..., y_n)$. *Furthermore let the distribution of* Y *be symbolized by* Q_Y .

Then

$$K := \{(y_1, ..., y_n) \in \mathbb{R}^n \mid \ L((y_1, ..., y_n); \cdot) \text{ has at least two maximizing points}\}$$

is a Borel subset of $\mathbb{R}^n$ *with* $\mathrm{Q}_Y(K) = 0$.

A Appendix

Lemma A.1 *Let* H_1, H_2 *be open subsets of* $\mathbb{R}^k$, *and let* $f : H_1 \times H_2 \to \mathbb{R}$ *be a* C^2-*mapping. Moreover there exist some* C^1-*diffeomorphism* g *from an open subset* U *of* $\mathbb{R}^k$ *onto an open subset* $g(U)$ *of* $\mathbb{R}^k$ *and a* C^1-*mapping* $\phi : H_2 \to \mathbb{R}^k$ *such that* $t - \phi(\gamma) \in U$ *and* $\nabla_\gamma f(t; \cdot) = g(t - \phi(\gamma))$ *hold for* $(t, \gamma) \in H_1 \times H_2$. *If* F *denotes a* C^2-*immersion from an open subset* A *of* $\mathbb{R}^p$ *into* H_2, *then* $\lambda^k(B_F) = 0$, *where* B_F *is defined to consist of all* $t \in H_1$ *such that* $\nabla_\alpha f(t, \cdot) \circ F = 0$ *and* $\det \nabla\nabla\alpha f(t, \cdot) \circ F = 0$ *for some* $\alpha \in H_2$.

Proof:

For $p = k$ we have

$$\nabla_\alpha f(t; \cdot) \circ F = 0, \ \det \nabla\nabla_\alpha f(t; \cdot) \circ F = 0 \quad \Leftrightarrow \quad \nabla_{F(\alpha)} f(t; \cdot) = 0, \ \det \nabla\nabla_{F(\alpha)} f(t; \cdot) = 0.$$

Thus, defining, $\tilde{\phi} : \tilde{U} \to \mathbf{R}^k, t \mapsto \phi(t) + g^{-1}(0)$, we may conclude $B_F \subseteq \tilde{\phi}(\{t \in \tilde{U} \mid \det J_t(\tilde{\phi}) = 0\})$. Note that g is a diffeomorphism with

$$\nabla_{F(\alpha)} f(t; \cdot) = g(t - \phi \circ F(\alpha)).$$

Sard's theorem (cf. appendix C, Proposition C.1) leads to $\lambda^k(B_F) = 0$.

Now let $p < k$.

Since F is a C^2-immersion, we may conclude from the rank theorem together with standard orthonormalization that for an arbitrary $\alpha \in A$ there are C^1- mappings $h^1_\alpha, ..., h^{k-p}_\alpha$ defined on an open neigbourhood U_α of α in A satisfying

$$h^i_\alpha(\tilde{\alpha})' J_{\tilde{\alpha}}(F) = 0, \quad h^i_\alpha(\tilde{\alpha})' h^i_\alpha(\tilde{\alpha}) = 1, \quad h^i_\alpha(\tilde{\alpha})' h^q_{\vartheta^j}(\tilde{\alpha}) = 0 \ (i \neq q)$$

for every $\tilde{\alpha} \in U_\alpha$ and arbitrary $i, q \in \{1, ..., k-p\}$.

By Lindelöf's theorem (cf. appendix C, Proposition C.2) we may select a sequence $(\alpha(s))_s$ in A such that $(U_\alpha(s))_s$ is a cover of A. For abbreviation we set $U_s := U_{\alpha(s)}, h_s^i := h_{\alpha(s)}^i$ $(i = 1, ..., k-p)$. Furthermore $x_\nu : \mathbb{R}^k \to \mathbb{R}$ denotes the projection on the $\nu-$th component $(\nu \in \{1, ..., k\})$.

We observe for $(t, \alpha) \in H_1 \times A$

$$\nabla_\alpha f(t; \cdot) \circ F = J_\alpha(F)' \nabla_{F(\alpha)} f(t; \cdot)$$

and

$$\begin{aligned} \nabla\nabla_\alpha f(t; \cdot) \circ F &= \sum_{\nu=1}^{k} \frac{\partial f(t; \cdot)}{\partial x_\nu}(F(\alpha)) \nabla\nabla_\alpha x_\nu \circ F + J_\alpha(F)' \nabla\nabla_{F(\alpha)} f(t; \cdot) J_\alpha(F) \\ &= \sum_{\nu=1}^{k} x_\nu \circ g(t - \phi \circ F(\alpha)) \nabla\nabla_\alpha x_\nu \circ F - J_\alpha(F)' J_{t-\phi\circ F(\alpha)}(g) J_\alpha(\phi \circ F). \end{aligned}$$

Let for $s \in \mathbb{N}$ the mappings $g^s, \phi^s : W_s \to \mathbb{R}^k$ be defined by $g^s(\alpha, b) := \sum_{i=1}^{k-p} b_i h_s^i(\alpha)$ and $\phi^s(\alpha, b) := \phi \circ F(\alpha) + g^{-1} \circ g^s(\alpha, b)$, where W_s consists of all $(\alpha, b) \in U_s \times \mathbb{R}^{k-p}$ with $\sum_{i=1}^{k-p} b_i h_s^i(\alpha) \in g(U)$. Notice that W_s is an open subset of $\mathbb{R}^k$, and that g^s is a C^1-mapping, which implies that ϕ^s is a C^1-mapping.

Now let $(\hat{t}, \hat{\alpha}) \in H_1 \times A$ with $\nabla_{\hat{\alpha}} f(\hat{t}; \cdot) \circ F = 0$. Then there exist some $s \in \mathbb{N}$ and a vector $\hat{b} := (\hat{b}_1, ..., \hat{b}_{k-p})$ from $\mathbb{R}^{k-p}$ with

$$g(\hat{t} - \phi \circ F(\hat{\alpha})) = \nabla_{F(\hat{\alpha})} f(\hat{t}; \cdot) = \sum_{i=1}^{k-p} \hat{b}_i h_s^i(\hat{\alpha}).$$

In particular $(\hat{\alpha}, \hat{b}) \in W_S$, and $g(\hat{t} - \phi \circ F(\hat{\alpha})) = g^s(\hat{\alpha}, \hat{b})$. Hence $\hat{t} = \phi^s(\hat{\alpha}, \hat{b})$.

As a consequence of $J_\alpha(F)' h_s^i(\alpha) = 0$ for $\alpha \in U_s$ as well as $i = 1, ..., k-p$ we obtain

$$\sum_{\nu=1}^{k} x_\nu \circ h_s^i(\alpha) \nabla\nabla_\alpha x_\nu \circ F = -J_\alpha(F)' J_\alpha(h_s^i),$$

and therefore

$$\nabla\nabla_{\hat{\alpha}} f(\hat{t}; \cdot) \circ F = -J_{\hat{\alpha}}(F)' J_{\hat{t}-\phi\circ F(\hat{\alpha})}(g) J_{\hat{\alpha}}(\phi^s(\cdot; \hat{b})).$$

Following the rules for determinants of partioned matrices (cf. [3], p.43, equation (II)), we get

$$\begin{aligned} |\det[J_{\hat\alpha}(F), h_s^1(\hat\alpha), ..., h_s^{k-p}(\hat\alpha)]' J_{\hat t - \phi\circ F(\hat\alpha)}(g) J_{(\hat\alpha,\hat b)}(\phi^s)| &= |\det J_{\hat\alpha}(F)' J_{\hat t - \phi\circ F(\hat\alpha)}(g) J_{\hat\alpha}(\phi^s(\cdot;\hat b))| \\ &= |\det \nabla\nabla_{\hat\alpha} f(\hat t;\cdot)\circ F| \end{aligned}$$

$[J_{\hat\alpha}(F), h_s^1(\hat\alpha), ..., h_s^{k-p}(\hat\alpha)]' J_{\hat t - \phi\circ F(\hat\alpha)}(g)$ has rank k since F is an immersion. Therefore we can conclude

$$\det \nabla\nabla_{\hat\alpha^j} f(\hat t;\cdot)\circ F^j = 0 \;\Leftrightarrow\; \det J_{(\hat\alpha,\hat b)}(\phi^s) = 0$$

Thus $B_F \subseteq N := \bigcup_{s=1}^{\infty} \phi^s(\{(\alpha,b)\in W_s \mid \det J_{(\alpha,b)}(\phi^s)=0\})$. Applying Sard's theorem (cf. appendix C, Proposition C.1), we get N as a set of Lebesgue-measure zero, which completes the proof. ■

B Appendix

Proof of Theorem 3.1:

Let us retake notations and assumptions from Theorem 3.1. Furthermore let us introduce for positive integer i,j the set M_{ij} consisting of all $y\in\mathcal{X}$ with $\nabla_{\vartheta^i} l(y,\cdot)\circ G^i = 0, \nabla_{\vartheta^j} l(y,\cdot)\circ G^j = 0$ and $l(y,G^i(\vartheta^i)) = l(y,G^j(\vartheta^j))$ for some ϑ^i,ϑ^j with $G^i(\vartheta^i)\neq G^j(\vartheta^j)$. Obviously, $K\subseteq \bigcup_{(i,j)\in\mathbb{N}\times\mathbb{N}} M_{ij}$. Therefore it remains to show

(1) M_{ij} is a λ^n-null set for arbitrary $i,j\in\mathbb{N}$.

(2) K is a Borel subset of $\mathcal{X}$.

proof of (1):

Let us define for $i,j\in\mathbb{N}$ the open subset $U := \{(\vartheta^i,\vartheta^j,y)\in\Theta^i\times\Theta^j\times\mathcal{X} \mid G^i(\vartheta^i)\neq G^j(\vartheta^j)\}$ of $\Theta^i\times\Theta^j\times\mathcal{X}$, and the set $Z := h^{-1}(\{0\})\cap\Theta^i\times\Theta^j\times(\mathbb{R}^n\setminus(L_i\cup L_j\cup L_{ij}))$, where

$$h: U\to\mathbf{R}^{r_i+r_j+1},\ (\vartheta^i,\vartheta^j,y)\mapsto \big(\nabla_{\vartheta^i} l(y;\cdot)\circ G^i, \nabla_{\vartheta^j} l(y;\cdot)\circ G^j, l(y;G^i(\vartheta^i)) - l(y;G^j(\vartheta^j))\big).$$

Let $(\hat{\vartheta}^i, \hat{\vartheta}^j, y)$ be an element of Z and $s \in \{1, ..., n\}$ with $\frac{\partial l(\cdot;G^i(\hat{\vartheta}^i))}{\partial y_s}(y) \neq \frac{\partial l(\cdot;G^j(\hat{\vartheta}^j))}{\partial y_s}(y)$. Using the Laplace-expansion of determinants and the rules of partioned matrices (cf. [3], p.43, equation (I)), we obtain

$$\begin{aligned} & \det[J_{(\hat{\vartheta}^i,\hat{\vartheta}^j)}(h(\cdot,\cdot,y)), \frac{\partial h}{\partial y_s}(\hat{\vartheta}^i, \hat{\vartheta}^j, y)] \\ = \; & \det(\nabla\nabla_{\hat{\vartheta}^i} l(y;\cdot) \circ G^i) \det(\nabla\nabla_{\hat{\vartheta}^j} l(y;\cdot) \circ G^j) \; \Big[\; \frac{\partial l(\cdot; G^i(\hat{\vartheta}^i))}{\partial y_s}(y) - \frac{\partial l(\cdot; G^j(\hat{\vartheta}^j))}{\partial y_s}(y) \; \Big] \neq 0, \end{aligned}$$

noticing $y \notin L_i \cup L_j$. Thus Z is a subset of $\tilde{U} := \{(\vartheta^i, \vartheta^j, y) \in U \mid \text{ rank of } J_{(\vartheta^i,\vartheta^j,y)}(h) = r_i + r_j + 1\}$, which is an open subset of $\mathbb{R}^{r_i+r_j+n}$. Therefore $h \mid \tilde{U}$ is a C^1-submersion and $\tilde{Z} := (h \mid \tilde{U})^{-1}(\{0\})$ is empty or a $(n-1)-$dim. C^1-submanifold of $\mathbf{R}^{r_i+r_j+n}$. We have

$$M_{ij} \setminus (L_i \cup L_j \cup L_{ij}) = Pr(Z) \subseteq Pr(\tilde{Z}),$$

where Pr denotes the canonical projection from $\mathbb{R}^{r_i+r_j+n}$ onto $\mathbb{R}^n$. By Sard's theorem (cf. appendix C, Proposition C.1) we get $\lambda^n(Pr(\tilde{Z})) = 0$ and, due to assumptions (3.1), (3.2),

$$\lambda^n(M_{ij}) = \lambda^n\big(M_{ij} \setminus (L_i \cup L_j \cup L_{ij})\big) \; + \lambda^n\big(L_i \cup L_j \cup L_{ij}\big) = 0.$$

proof of (2):

For positive integers i,j we define the open subset $N_{ij} := \{(\vartheta^i, \vartheta^j) \in \Theta^i \times \Theta^j \mid G^i(\vartheta^i) \neq G^j(\vartheta^j)\}$ of $\Theta^i \times \Theta^j$ respectively $\mathbf{R}^{r_i+r_j}$. Applying Lindelöf's Theorem (cf. appendix C, Proposition C.2) and using the local compactness of N_{ij} we get a sequence $(U_\nu \times V_\nu)_\nu$ of compact sets $U_\nu \times V_\nu$ satisfying $N_{ij} = \bigcup_{\nu=1}^{\infty} U_\nu \times V_\nu$. Since each set $U_\nu \times V_\nu$ is sequentially compact, every $y \in \mathcal{X}$ being an accumulation point of the set

$$G_\nu^{ij} := \big\{y \in \mathcal{X} \mid l(y; G^i(\vartheta^i)) = l(y; G^j(\vartheta^j)) = \max_{\vartheta \in \vartheta} l_G(y; \vartheta) \text{ for some } (\vartheta^i, \vartheta^j) \in U_\nu \times V_\nu\big\}$$

is also an element of G_ν^{ij}, i.e. the sets G_ν^{ij} are closed in $\mathcal{X}$ and therefore Borel subsets. Then K is a Borel subset because

$$K = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcup_{\nu=1}^{\infty} G_\nu^{ij}$$

Therefore (2) is shown, which completes the proof. ■

C Appendix

Proposition C.1 (Special versions of Sard's theorem) *Let $f : U \to \mathbb{R}^t$ denote a C^r−mapping ($r \in \mathbb{N}$) from an open subset U of $\mathbb{R}^s$ into $\mathbb{R}^t$. Then we can state:*

.1 If $s = t$, then $f(\{x \in U \mid \det\ J_x(f) = 0\})$ is a λ^t−null set.

.2 If $A \subseteq U$ is a m−dimensional C^r−submanifold of $\mathbb{R}^s$ with $0 \leq m < t$, then $f(A)$ is a λ^t−null set.

These versions of Sard's theorem may be found in [1] (Corollary 3.3.17.4 and Theorem 4.3.1) or in [4] (p. 69) or in [10] (remark on Definition 3.3 and Theorem 3.1).

Proposition C.2 (Lindelöf's theorem) *Let (Ω, τ) denote a topological Hausdorff space such that τ has a countable base. Then every open cover of a subset $A \subseteq \Omega$ has a countable subcover.*

For a proof cf. e.g. [6] (p.49).

References

[1] Berger, M. and Gostiaux, B., *Differential Geometry: Manifolds, Curves and Surfaces*, Springer, New York, 1988.

[2] Brown, L., *Fundamentals of Statistical Exponential Families*, IMS Monograph Series, Hayward, California, 1986.

[3] Gantmacher, F.R., *Matrizenrechnung I*, Deutscher Verlag der Wissenschaften, Berlin (Ost) (1958).

[4] Hirsch, M.W., *Differential Topology*, Springer, New York, 1976.

[5] Johnson, N.L., Kotz, S. and Balakrishnan, N., *Continuous Univariate Distributions Volume 2*, Wiley, New York, 1995.

[6] Kelley, J.F., *General Topology*, van Nostrand, New York, 1955.

[7] Pazman, A., *Nonlinear Least Squares - Uniqueness versus Ambiguity*, Mathematische Operationsforschung, series Statistics, 15 (1984), 323-336.

[8] Pazman, A., *On the Uniqueness of the M.L. Estimates in Curved Exponential Families*, Kybernetika 22 (1986), 124-132.

[9] Pazman, A., *Nonlinear Statistical Models*, Kluwer, Dordrecht, 1993.

[10] Sternberg, S. *Lectures on Differential Geometry*, Prentice Hall, Englewood Cliffs, 1964.

[11] Witting, H., *Mathematische Statistik I*, Teubner, Stuttgart, 1985.

SFB 649 Discussion Paper Series 2006

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Calibration Risk for Exotic Options" by Kai Detlefsen and Wolfgang K. Härdle, January 2006.
002 "Calibration Design of Implied Volatility Surfaces" by Kai Detlefsen and Wolfgang K. Härdle, January 2006.
003 "On the Appropriateness of Inappropriate VaR Models" by Wolfgang Härdle, Zdeněk Hlávka and Gerhard Stahl, January 2006.
004 "Regional Labor Markets, Network Externalities and Migration: The Case of German Reunification" by Harald Uhlig, January/February 2006.
005 "British Interest Rate Convergence between the US and Europe: A Recursive Cointegration Analysis" by Enzo Weber, January 2006.
006 "A Combined Approach for Segment-Specific Analysis of Market Basket Data" by Yasemin Boztuğ and Thomas Reutterer, January 2006.
007 "Robust utility maximization in a stochastic factor model" by Daniel Hernández–Hernández and Alexander Schied, January 2006.
008 "Economic Growth of Agglomerations and Geographic Concentration of Industries - Evidence for Germany" by Kurt Geppert, Martin Gornig and Axel Werwatz, January 2006.
009 "Institutions, Bargaining Power and Labor Shares" by Benjamin Bental and Dominique Demougin, January 2006.
010 "Common Functional Principal Components" by Michal Benko, Wolfgang Härdle and Alois Kneip, Jauary 2006.
011 "VAR Modeling for Dynamic Semiparametric Factors of Volatility Strings" by Ralf Brüggemann, Wolfgang Härdle, Julius Mungo and Carsten Trenkler, February 2006.
012 "Bootstrapping Systems Cointegration Tests with a Prior Adjustment for Deterministic Terms" by Carsten Trenkler, February 2006.
013 "Penalties and Optimality in Financial Contracts: Taking Stock" by Michel A. Robe, Eva-Maria Steiger and Pierre-Armand Michel, February 2006.
014 "Core Labour Standards and FDI: Friends or Foes? The Case of Child Labour" by Sebastian Braun, February 2006.
015 "Graphical Data Representation in Bankruptcy Analysis" by Wolfgang Härdle, Rouslan Moro and Dorothea Schäfer, February 2006.
016 "Fiscal Policy Effects in the European Union" by Andreas Thams, February 2006.
017 "Estimation with the Nested Logit Model: Specifications and Software Particularities" by Nadja Silberhorn, Yasemin Boztuğ and Lutz Hildebrandt, March 2006.
018 "The Bologna Process: How student mobility affects multi-cultural skills and educational quality" by Lydia Mechtenberg and Roland Strausz, March 2006.
019 "Cheap Talk in the Classroom" by Lydia Mechtenberg, March 2006.
020 "Time Dependent Relative Risk Aversion" by Enzo Giacomini, Michael Handel and Wolfgang Härdle, March 2006.
021 "Finite Sample Properties of Impulse Response Intervals in SVECMs with Long-Run Identifying Restrictions" by Ralf Brüggemann, March 2006.
022 "Barrier Option Hedging under Constraints: A Viscosity Approach" by Imen Bentahar and Bruno Bouchard, March 2006.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

023 "How Far Are We From The Slippery Slope? The Laffer Curve Revisited" by Mathias Trabandt and Harald Uhlig, April 2006.
024 "e-Learning Statistics – A Selective Review" by Wolfgang Härdle, Sigbert Klinke and Uwe Ziegenhagen, April 2006.
025 "Macroeconomic Regime Switches and Speculative Attacks" by Bartosz Maćkowiak, April 2006.
026 "External Shocks, U.S. Monetary Policy and Macroeconomic Fluctuations in Emerging Markets" by Bartosz Maćkowiak, April 2006.
027 "Institutional Competition, Political Process and Holdup" by Bruno Deffains and Dominique Demougin, April 2006.
028 "Technological Choice under Organizational Diseconomies of Scale" by Dominique Demougin and Anja Schöttner, April 2006.
029 "Tail Conditional Expectation for vector-valued Risks" by Imen Bentahar, April 2006.
030 "Approximate Solutions to Dynamic Models – Linear Methods" by Harald Uhlig, April 2006.
031 "Exploratory Graphics of a Financial Dataset" by Antony Unwin, Martin Theus and Wolfgang Härdle, April 2006.
032 "When did the 2001 recession *really* start?" by Jörg Polzehl, Vladimir Spokoiny and Cătălin Stărică, April 2006.
033 "Varying coefficient GARCH versus local constant volatility modeling. Comparison of the predictive power" by Jörg Polzehl and Vladimir Spokoiny, April 2006.
034 "Spectral calibration of exponential Lévy Models [1]" by Denis Belomestny and Markus Reiß, April 2006.
035 "Spectral calibration of exponential Lévy Models [2]" by Denis Belomestny and Markus Reiß, April 2006.
036 "Spatial aggregation of local likelihood estimates with applications to classification" by Denis Belomestny and Vladimir Spokoiny, April 2006.
037 "A jump-diffusion Libor model and its robust calibration" by Denis Belomestny and John Schoenmakers, April 2006.
038 "Adaptive Simulation Algorithms for Pricing American and Bermudan Options by Local Analysis of Financial Market" by Denis Belomestny and Grigori N. Milstein, April 2006.
039 "Macroeconomic Integration in Asia Pacific: Common Stochastic Trends and Business Cycle Coherence" by Enzo Weber, May 2006.
040 "In Search of Non-Gaussian Components of a High-Dimensional Distribution" by Gilles Blanchard, Motoaki Kawanabe, Masashi Sugiyama, Vladimir Spokoiny and Klaus-Robert Müller, May 2006.
041 "Forward and reverse representations for Markov chains" by Grigori N. Milstein, John G. M. Schoenmakers and Vladimir Spokoiny, May 2006.
042 "Discussion of 'The Source of Historical Economic Fluctuations: An Analysis using Long-Run Restrictions' by Neville Francis and Valerie A. Ramey" by Harald Uhlig, May 2006.
043 "An Iteration Procedure for Solving Integral Equations Related to Optimal Stopping Problems" by Denis Belomestny and Pavel V. Gapeev, May 2006.
044 "East Germany's Wage Gap: A non-parametric decomposition based on establishment characteristics" by Bernd Görzig, Martin Gornig and Axel Werwatz, May 2006.
045 "Firm Specific Wage Spread in Germany - Decomposition of regional differences in inter firm wage dispersion" by Bernd Görzig, Martin Gornig and Axel Werwatz, May 2006.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

046 "Produktdiversifizierung: Haben die ostdeutschen Unternehmen den Anschluss an den Westen geschafft? – Eine vergleichende Analyse mit Mikrodaten der amtlichen Statistik" by Bernd Görzig, Martin Gornig and Axel Werwatz, May 2006.

047 "The Division of Ownership in New Ventures" by Dominique Demougin and Oliver Fabel, June 2006.

048 "The Anglo-German Industrial Productivity Paradox, 1895-1938: A Restatement and a Possible Resolution" by Albrecht Ritschl, May 2006.

049 "The Influence of Information Costs on the Integration of Financial Markets: Northern Europe, 1350-1560" by Oliver Volckart, May 2006.

050 "Robust Econometrics" by Pavel Čížek and Wolfgang Härdle, June 2006.

051 "Regression methods in pricing American and Bermudan options using consumption processes" by Denis Belomestny, Grigori N. Milstein and Vladimir Spokoiny, July 2006.

052 "Forecasting the Term Structure of Variance Swaps" by Kai Detlefsen and Wolfgang Härdle, July 2006.

053 "Governance: Who Controls Matters" by Bruno Deffains and Dominique Demougin, July 2006.

054 "On the Coexistence of Banks and Markets" by Hans Gersbach and Harald Uhlig, August 2006.

055 "Reassessing Intergenerational Mobility in Germany and the United States: The Impact of Differences in Lifecycle Earnings Patterns" by Thorsten Vogel, September 2006.

056 "The Euro and the Transatlantic Capital Market Leadership: A Recursive Cointegration Analysis" by Enzo Weber, September 2006.

057 "Discounted Optimal Stopping for Maxima in Diffusion Models with Finite Horizon" by Pavel V. Gapeev, September 2006.

058 "Perpetual Barrier Options in Jump-Diffusion Models" by Pavel V. Gapeev, September 2006.

059 "Discounted Optimal Stopping for Maxima of some Jump-Diffusion Processes" by Pavel V. Gapeev, September 2006.

060 "On Maximal Inequalities for some Jump Processes" by Pavel V. Gapeev, September 2006.

061 "A Control Approach to Robust Utility Maximization with Logarithmic Utility and Time-Consistent Penalties" by Daniel Hernández–Hernández and Alexander Schied, September 2006.

062 "On the Difficulty to Design Arabic E-learning System in Statistics" by Taleb Ahmad, Wolfgang Härdle and Julius Mungo, September 2006.

063 "Robust Optimization of Consumption with Random Endowment" by Wiebke Wittmüß, September 2006.

064 "Common and Uncommon Sources of Growth in Asia Pacific" by Enzo Weber, September 2006.

065 "Forecasting Euro-Area Variables with German Pre-EMU Data" by Ralf Brüggemann, Helmut Lütkepohl and Massimiliano Marcellino, September 2006.

066 "Pension Systems and the Allocation of Macroeconomic Risk" by Lans Bovenberg and Harald Uhlig, September 2006.

067 "Testing for the Cointegrating Rank of a VAR Process with Level Shift and Trend Break" by Carsten Trenkler, Pentti Saikkonen and Helmut Lütkepohl, September 2006.

068 "Integral Options in Models with Jumps" by Pavel V. Gapeev, September 2006.

069 "Constrained General Regression in Pseudo-Sobolev Spaces with Application to Option Pricing" by Zdeněk Hlávka and Michal Pešta, September 2006.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

070 "The Welfare Enhancing Effects of a Selfish Government in the Presence of Uninsurable, Idiosyncratic Risk" by R. Anton Braun and Harald Uhlig, September 2006.
071 "Color Harmonization in Car Manufacturing Process" by Anton Andriyashin, Michal Benko, Wolfgang Härdle, Roman Timofeev and Uwe Ziegenhagen, October 2006.
072 "Optimal Interest Rate Stabilization in a Basic Sticky-Price Model" by Matthias Paustian and Christian Stoltenberg, October 2006.
073 "Real Balance Effects, Timing and Equilibrium Determination" by Christian Stoltenberg, October 2006.
074 "Multiple Disorder Problems for Wiener and Compound Poisson Processes With Exponential Jumps" by Pavel V. Gapeev, October 2006.
075 "Inhomogeneous Dependency Modelling with Time Varying Copulae" by Enzo Giacomini, Wolfgang K. Härdle, Ekaterina Ignatieva and Vladimir Spokoiny, November 2006.
076 "Convenience Yields for CO_2 Emission Allowance Futures Contracts" by Szymon Borak, Wolfgang Härdle, Stefan Trück and Rafal Weron, November 2006.
077 "Estimation of Default Probabilities with Support Vector Machines" by Shiyi Chen, Wolfgang Härdle and Rouslan Moro, November 2006.
078 "GHICA - Risk Analysis with GH Distributions and Independent Components" by Ying Chen, Wolfgang Härdle and Vladimir Spokoiny, November 2006.
079 "Do Individuals Recognize Cascade Behavior of Others? - An Experimental Study –" by Tim Grebe, Julia Schmid and Andreas Stiehler, November 2006.
080 "The Uniqueness of Extremum Estimation" by Volker Krätschmer, December 2006.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".